

1/69

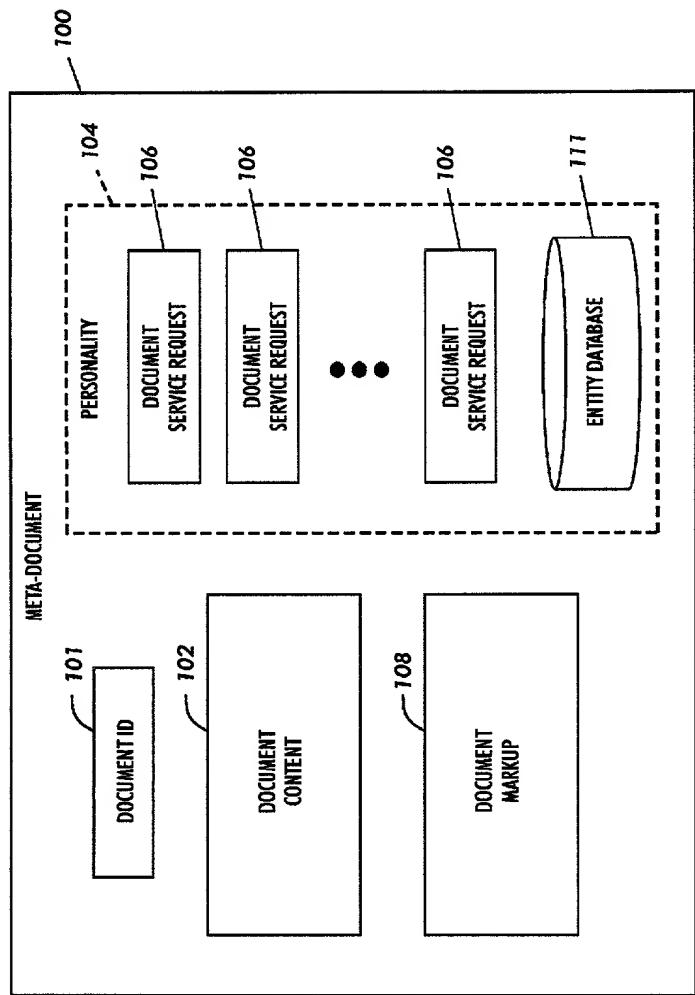
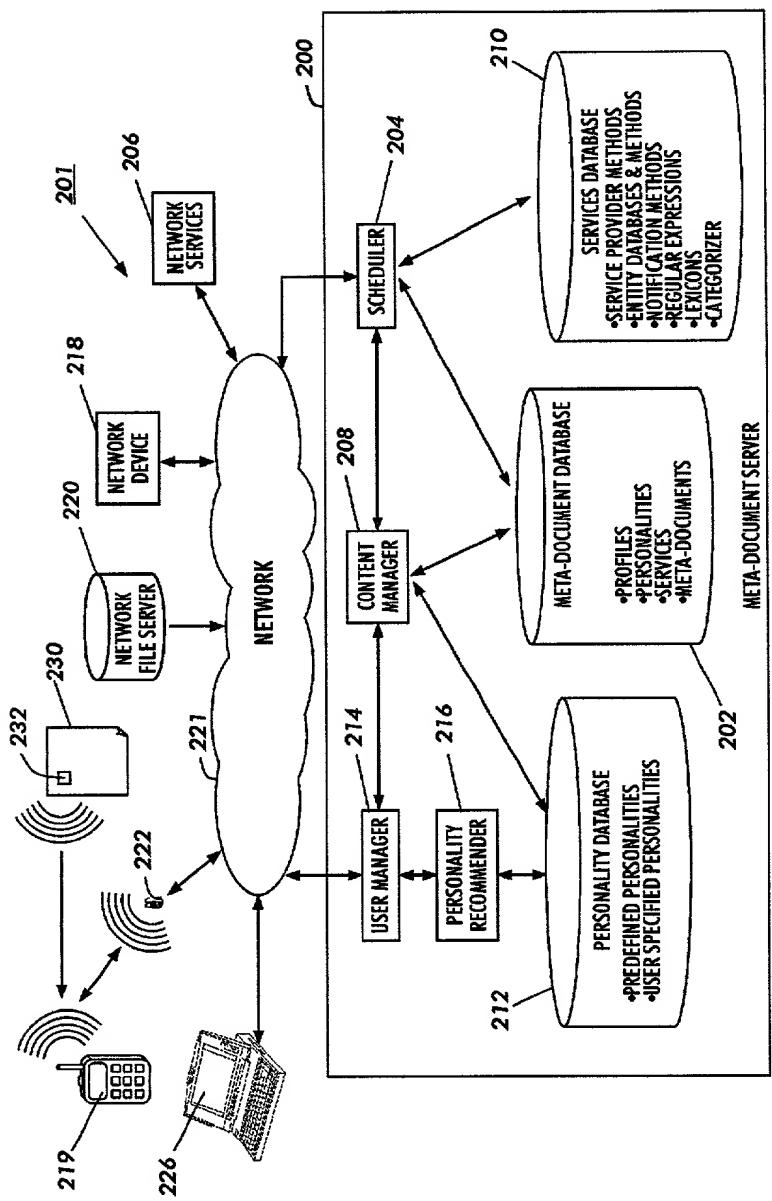


FIG. 1

2/69



3/69

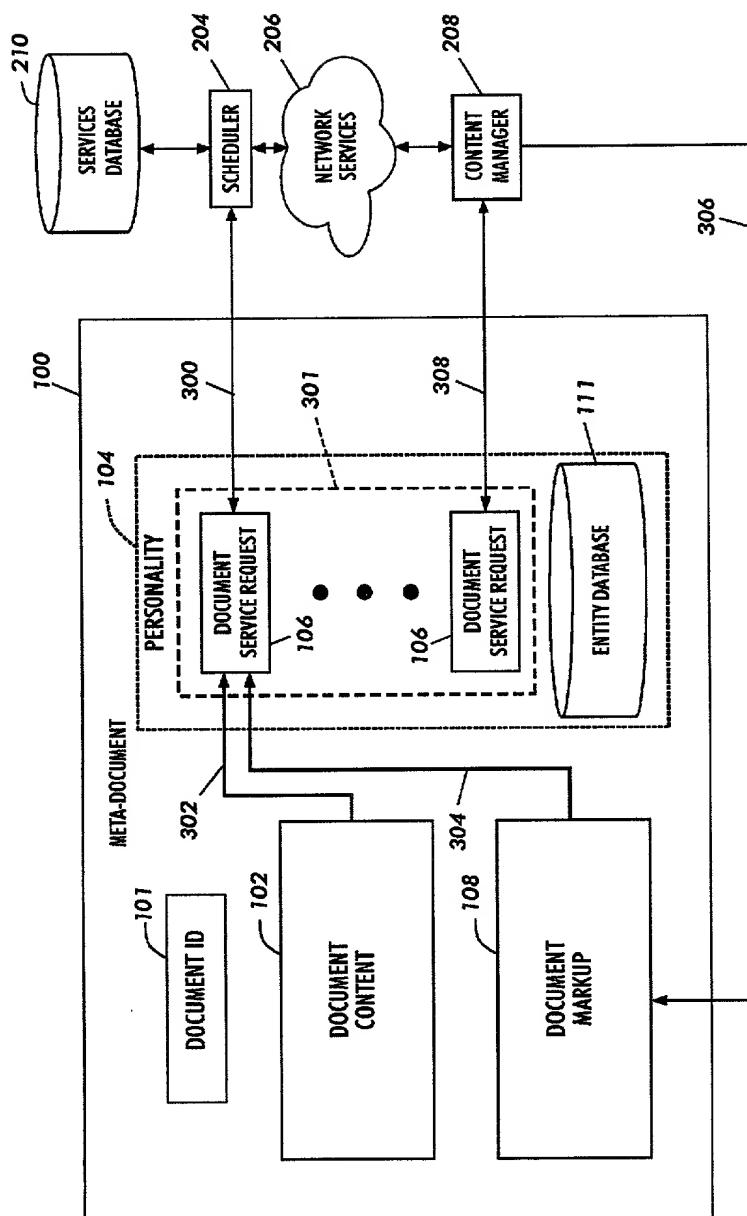


FIG. 3

4/69

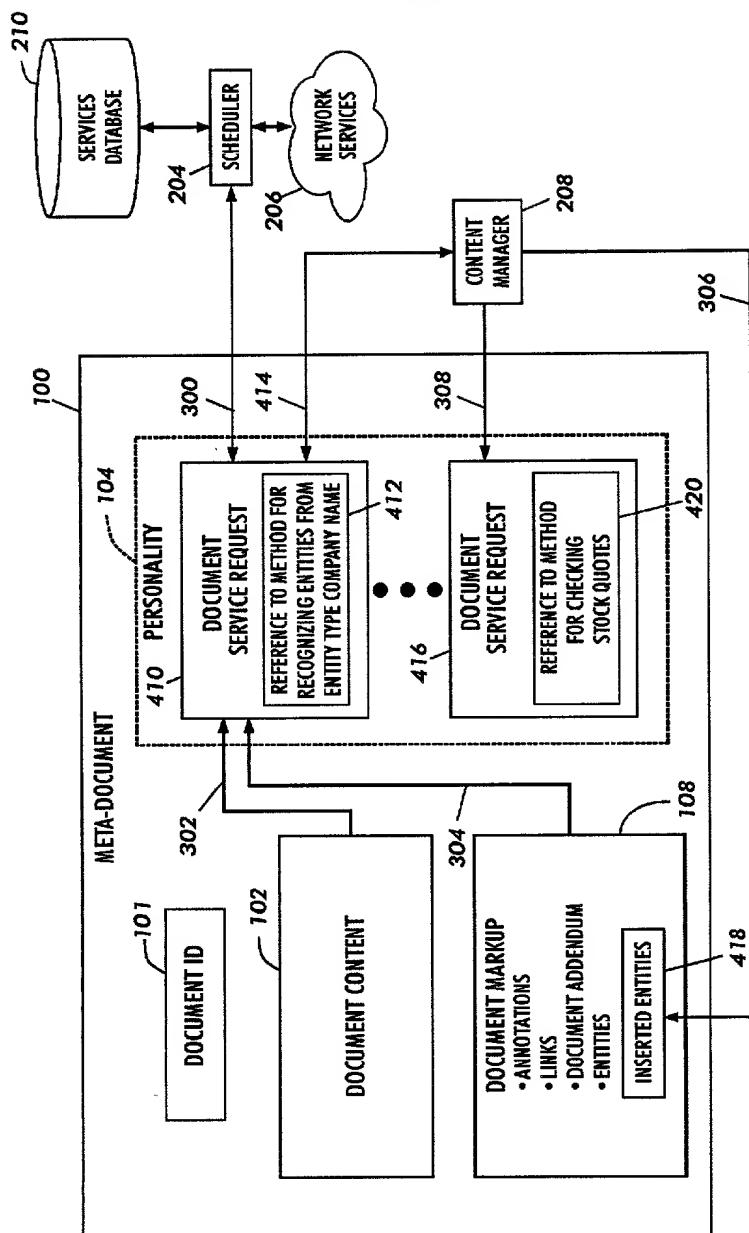


FIG. 4

5/69

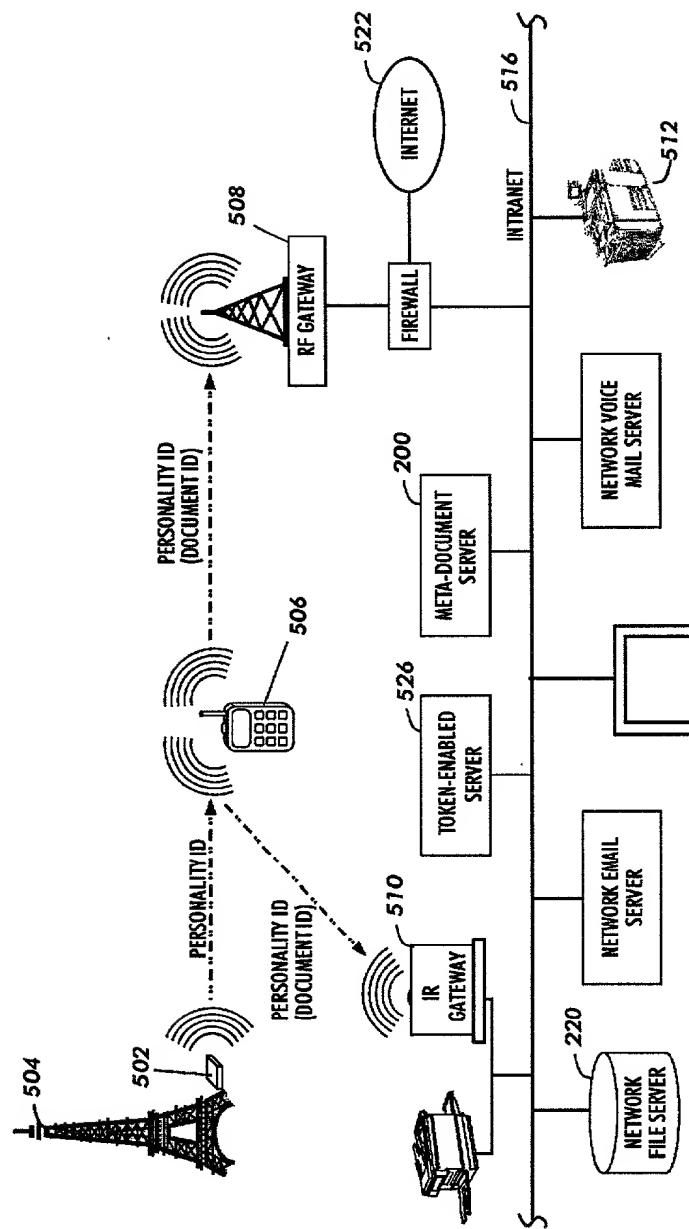


FIG. 5

6/69

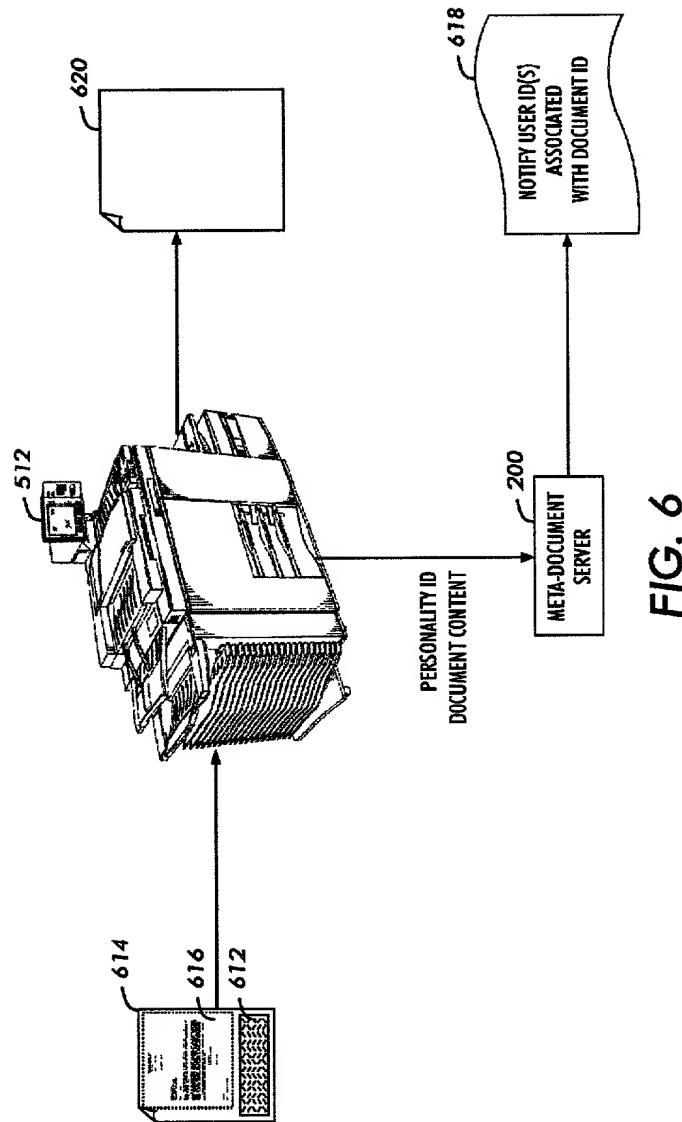


FIG. 6

7/69

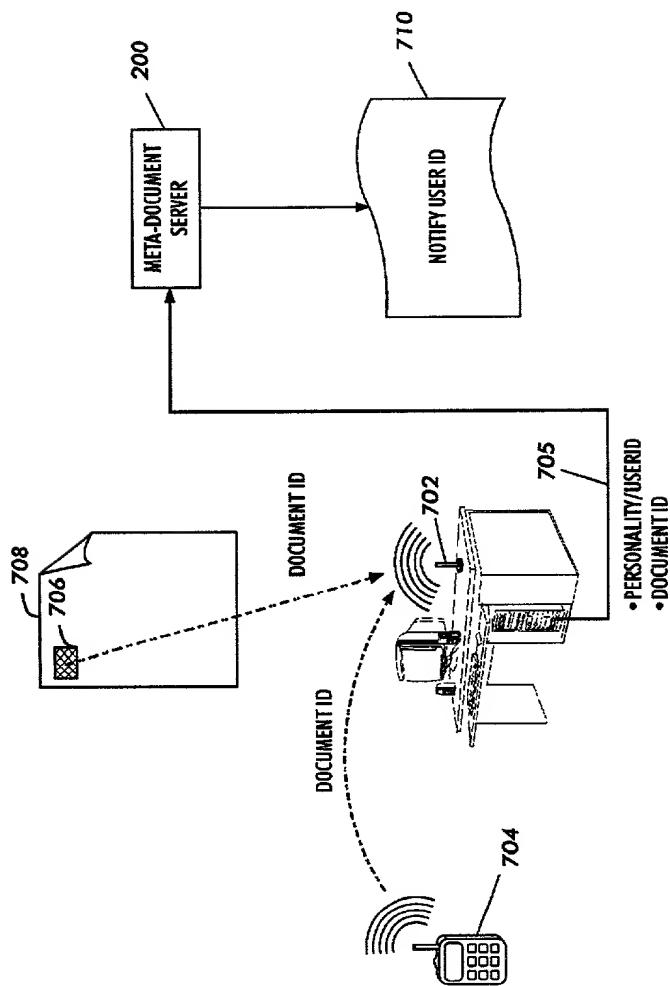
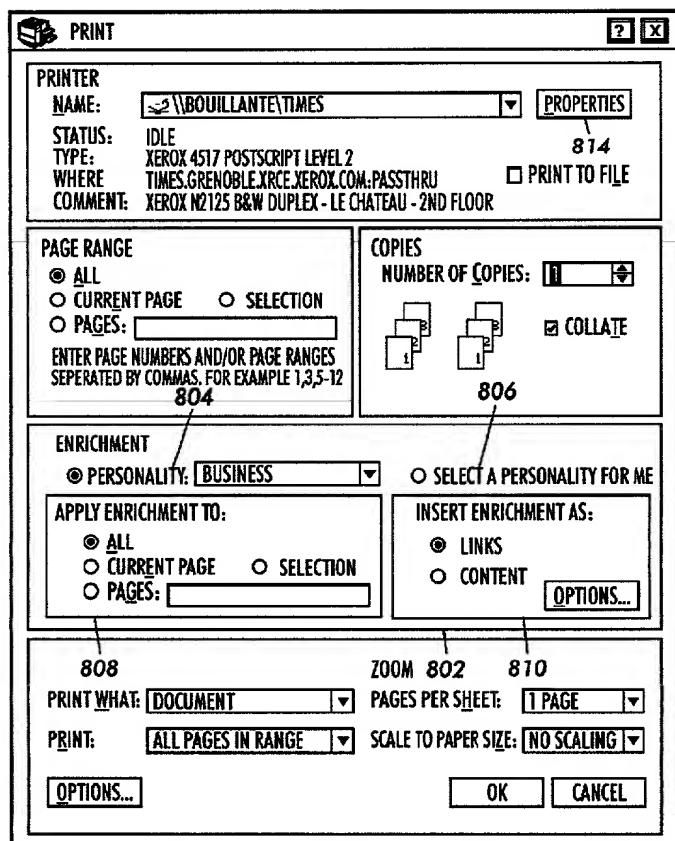


FIG. 7

8/69



800

FIG. 8

9/69

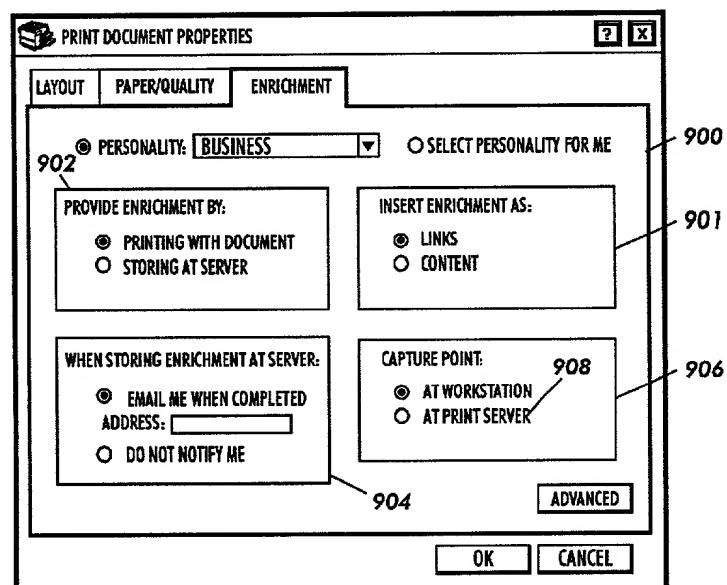


FIG. 9

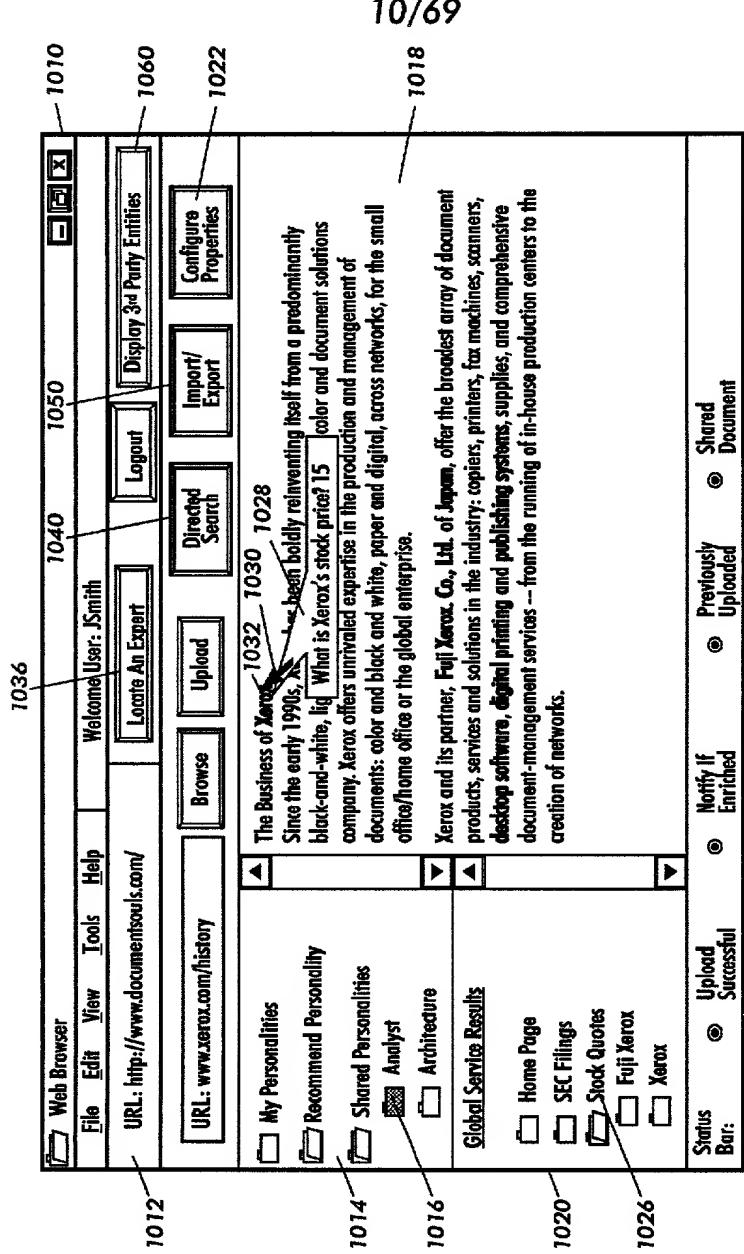


FIG. 10

11/69

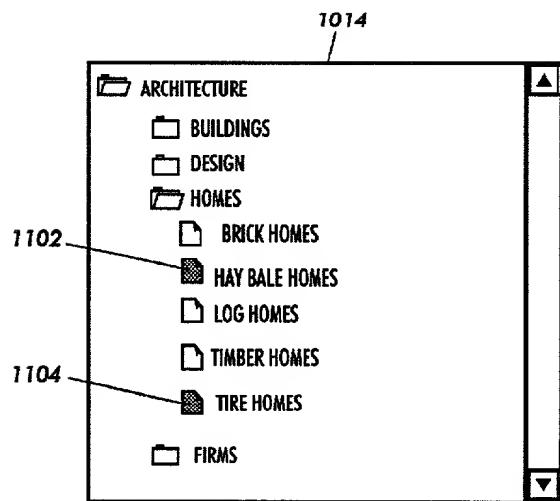


FIG. 11

12/69

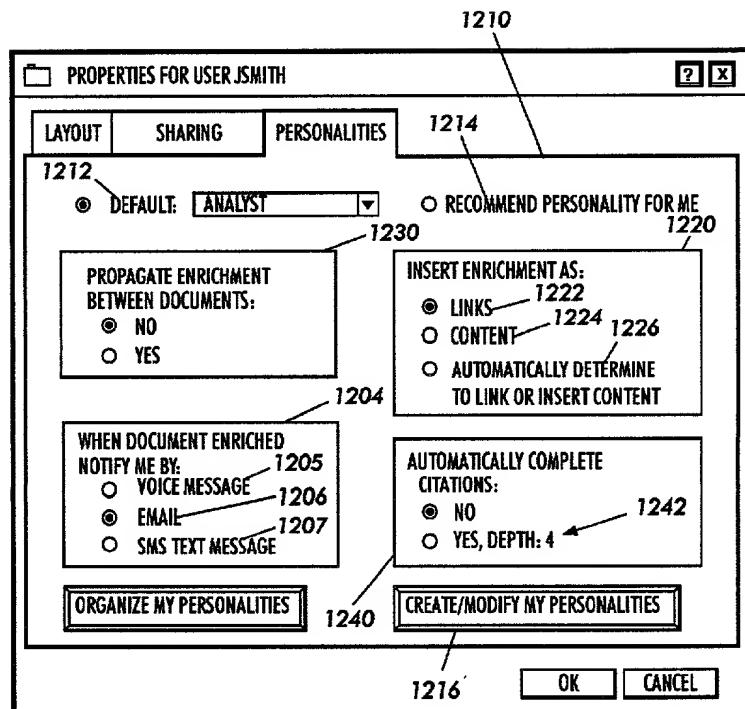


FIG. 12

13/69

CREATE/MODIFY A GROUP OF EXISTING PERSONALITIES													
<input type="checkbox"/> LIST OF AVAILABLE PERSONALITIES	<table border="1"> <thead> <tr> <th>Name</th> <th>Creation Date</th> </tr> </thead> <tbody> <tr> <td>Technology Watch in Computer Sciences</td> <td>2001-03-16 14:08:06</td> </tr> <tr> <td>Watch Business</td> <td>2001-03-16 14:08:08</td> </tr> <tr> <td>Watch Legal</td> <td>2001-03-16 14:08:10</td> </tr> <tr> <td>Watch Intellectual Property</td> <td>2001-03-16 14:08:10</td> </tr> <tr> <td>Scientific Watch in Medicine</td> <td>2001-03-16 14:08:10</td> </tr> </tbody> </table>	Name	Creation Date	Technology Watch in Computer Sciences	2001-03-16 14:08:06	Watch Business	2001-03-16 14:08:08	Watch Legal	2001-03-16 14:08:10	Watch Intellectual Property	2001-03-16 14:08:10	Scientific Watch in Medicine	2001-03-16 14:08:10
Name	Creation Date												
Technology Watch in Computer Sciences	2001-03-16 14:08:06												
Watch Business	2001-03-16 14:08:08												
Watch Legal	2001-03-16 14:08:10												
Watch Intellectual Property	2001-03-16 14:08:10												
Scientific Watch in Medicine	2001-03-16 14:08:10												
<input type="checkbox"/> LIST OF ALL WATCH BUSINESS SERVICES	<p>those Services are classified by Service Categories, a ticked checkbox meaning that the according service is attached to the selected personality.</p> <p>Click on this icon to view all available service categories</p> <ul style="list-style-type: none"> Search Engine <input type="checkbox"/> * <i>luckygoat People Google Lucky Result</i> Stock Quotes <input type="checkbox"/> * <i>bloomberg Stock Quotes</i> 												

FIG. 13

14/69

	NAME	ENTITY TYPE	ANNOTATE	DEPTH	PROVIDER	COST	QUERY	REFRESH
SEL	COMPUTING	COMPUTING	ALWAYS	0	YAHOO	FREE	LINK	DAILY
	BIOLOGY	BIOLOGY	ALWAYS	0	GOOGLELUCKY	NULL	LINK	NEVER
	MEDICINE	MEDICINE	ALWAYS	0	ALTAVISTA	FEF	LINK	AUTO
	APPLY PAIRS OF ENTITIES AND IDENTIFY 3RD PARTY ENTITIES	COMPANIES	EXPERT	0	GOOGLELUCKY	NULL	CONTENT	NULL
	APPLY PAIRS OF ENTITIES USING GOOGLELUCKY	PRODUCTS	ALWAYS	2	NORTHERNLIGHTS	NULL	LINK	WEEKLY
	COUNTRIES	COUNTRIES	NOVICE	0	GOOGLELUCKY	FEF	CONTENT	UNTIL 2003
	ASTRONOMY	ASTRONOMY	ALWAYS	1	GOOGLE	NULL	LINK	MONTHLY
	PEOPLE	PEOPLE	ALWAYS	3	GOOGLE PEOPLE	NULL	LINK	NULL

◀ ▶

1424

1432

VIEW HIERARCHY

SET DEPTH

1378 1400 1426 1412 1414 1430 1442 1440 1444 1450

1404 1422 1420 1408

FIG. 14

15/69

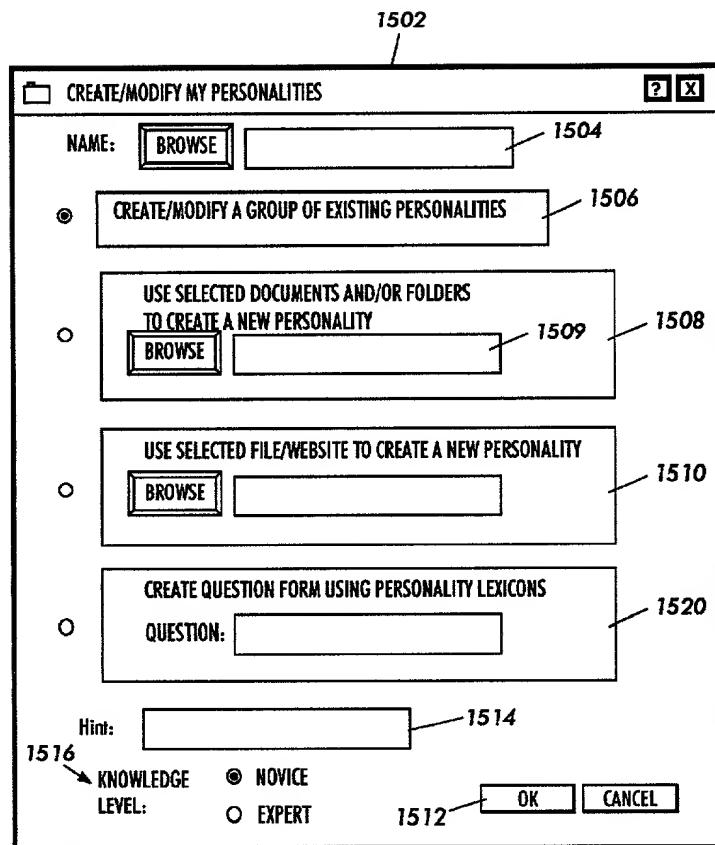


FIG. 15

16/69

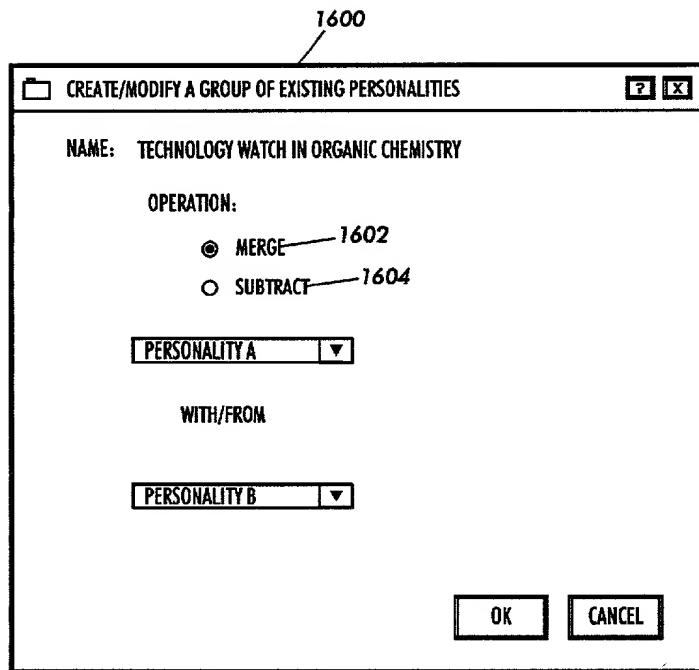
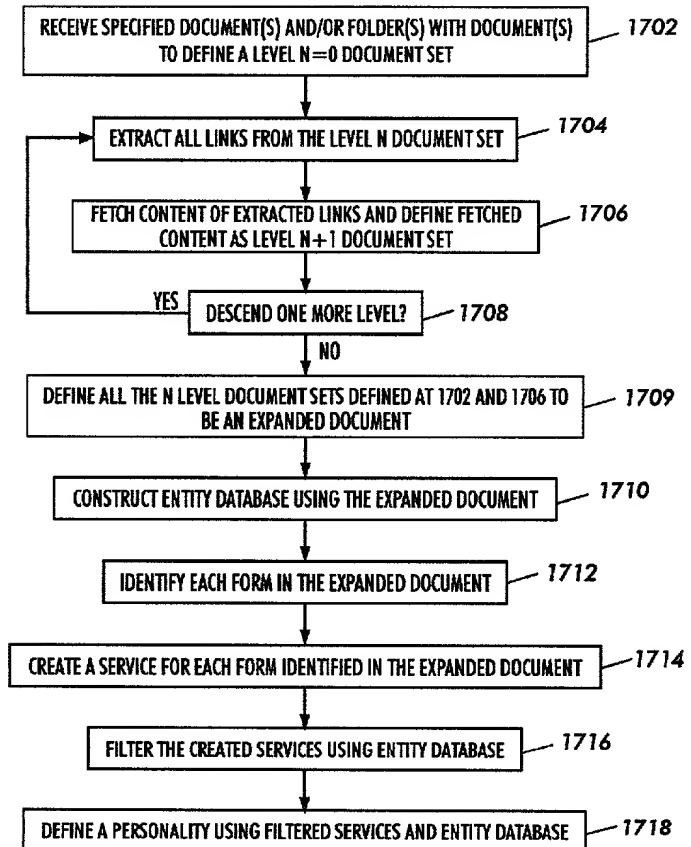


FIG. 16

17/69

**FIG. 17**

18/69

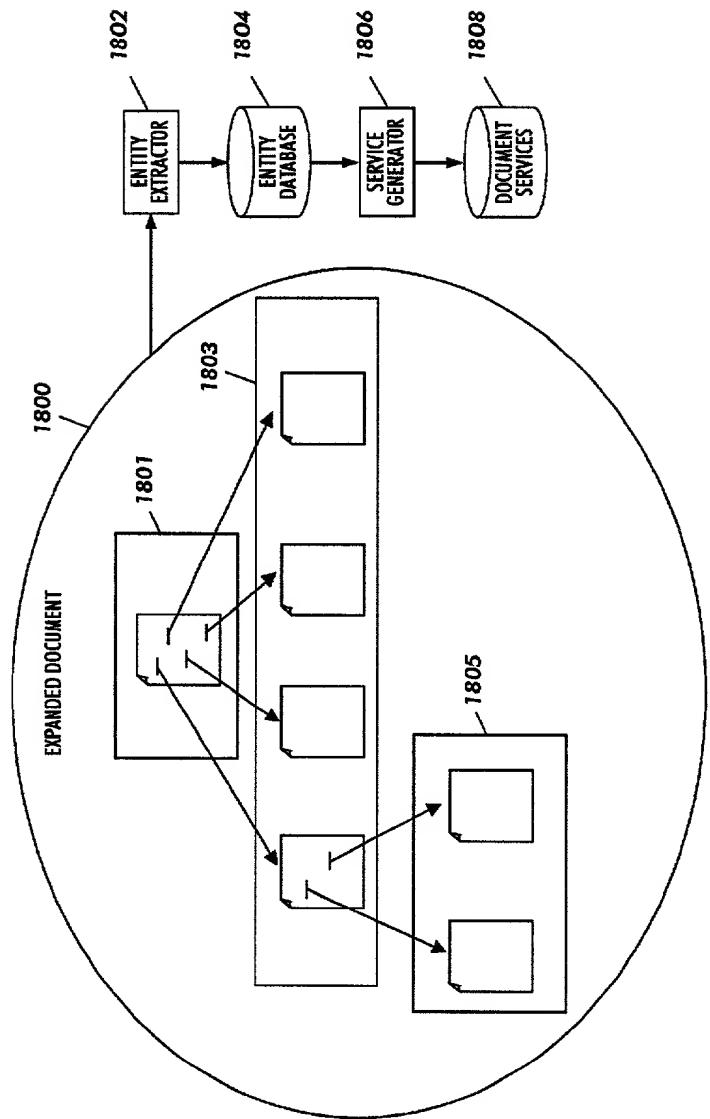


FIG. 18

19/69

1900

SCIENTIFIC CONTENT PROVIDER (S.C.P.):	
1901	
SEARCH	
SEARCH IN:	
1902	<input type="radio"/> ALL FOLDERS
1903	<input type="radio"/> PHYSICS FOLDER
1904	<input type="radio"/> CHEMISTRY FOLDER
1905	<input type="radio"/> BIOLOGY FOLDER

FIG. 19

20/69

SCIENTIFIC SEARCH □ ▶						
SEL	NAME	ENTITY TYPE	ANNOTATE	DEPTH	PROVIDER	COST
<input checked="" type="checkbox"/>	ALL FOLDERS RESULT OF SCIENTIFIC CONTENT PROVIDER	STANDARD	ALWAYS	0	S.C.P.	FREE
<input type="checkbox"/>	PHYSICS RESULT OF SCIENTIFIC CONTENT PROVIDER	PHYSICS	ALWAYS	0	S.C.P.	FREE
<input type="checkbox"/>	CHEMISTRY RESULT OF SCIENTIFIC CONTENT PROVIDER	CHEMISTRY	ALWAYS	0	S.C.P.	FREE
<input type="checkbox"/>	BIOLOGY RESULT OF SCIENTIFIC CONTENT PROVIDER	BIOLOGY	ALWAYS	0	S.C.P.	FREE

2002
2004

FIG. 20

21/69

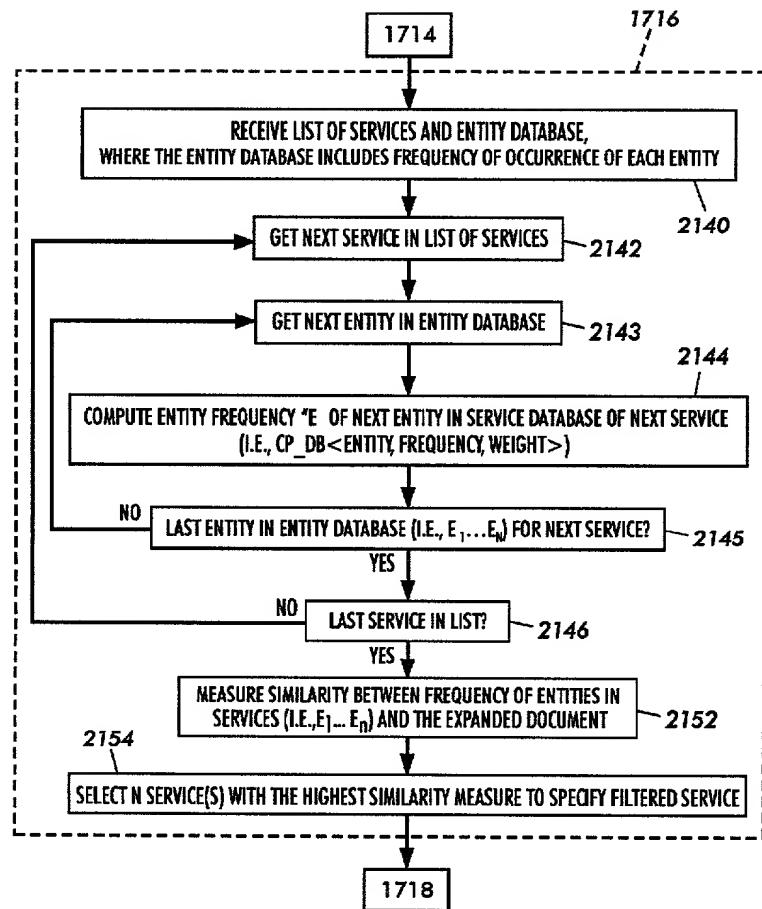


FIG. 21

22/69

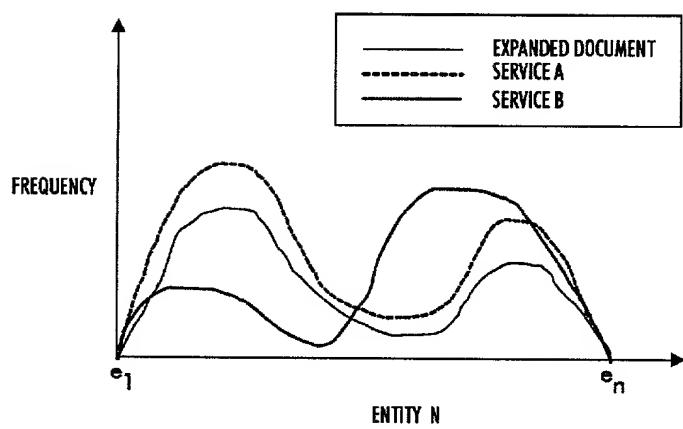


FIG. 22

23/69

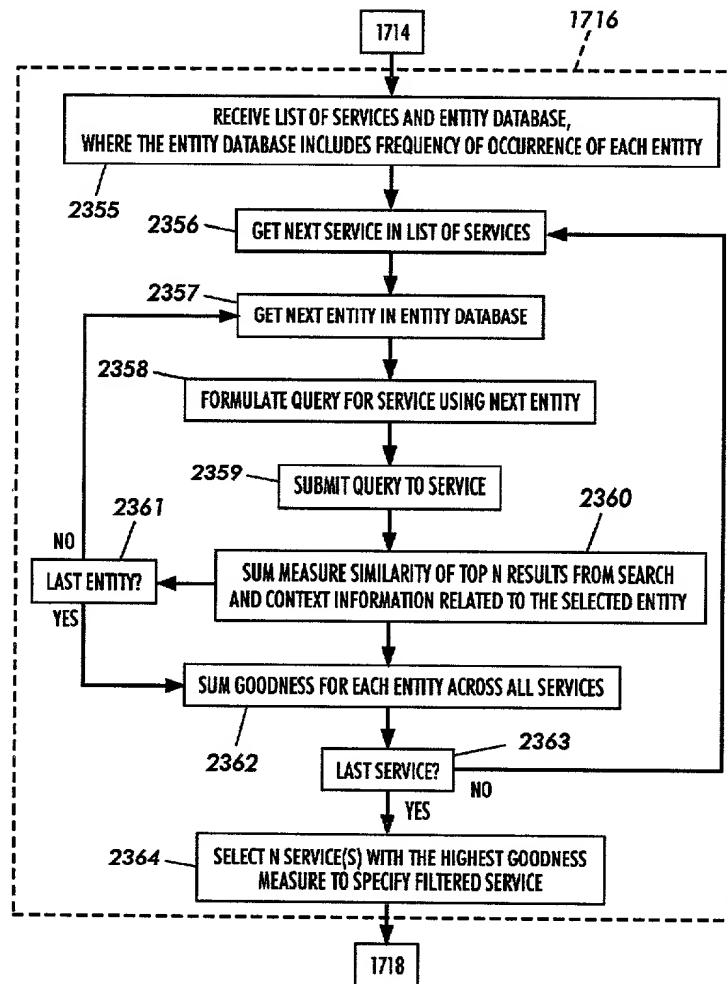


FIG. 23

24/69

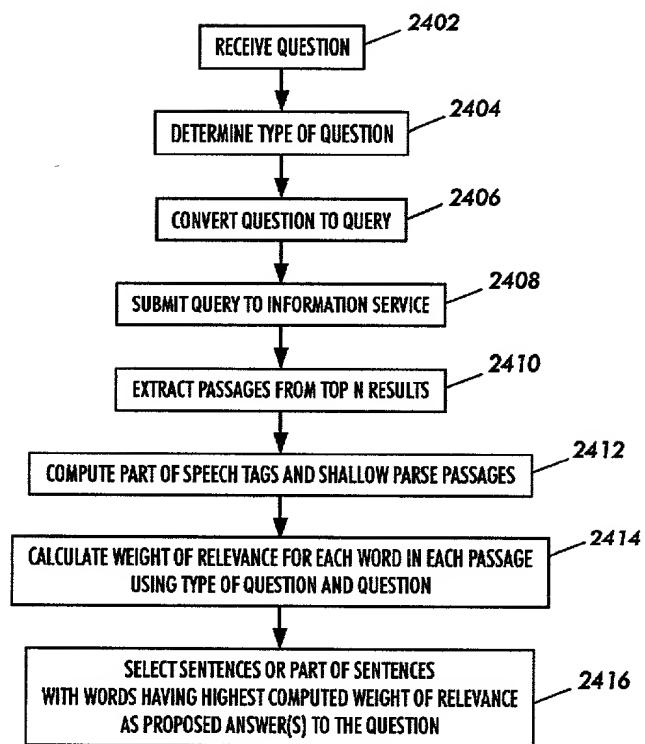


FIG. 24

25/69

E-LEARNING SERVICES ASSOCIATED WITH E-LEARNING PERSONALITY	
	SEL
2504	<input checked="" type="checkbox"/> LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR DEFINITIONS
2506	<input checked="" type="checkbox"/> LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR TRANSLATIONS
2508	<input type="checkbox"/> LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO A TUTORIAL CONCERNING EACH TEXT UNIT
2510	<input type="checkbox"/> LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO AN EXPERT CONCERNING IDENTIFIED SUBJECTS
2512	<input type="checkbox"/> LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO AN INTERACTIVE COURSE ABOUT IDENTIFIED SUBJECTS
2514	<input checked="" type="checkbox"/> LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO A LIST OF REAL WORLD CLASSES ABOUT IDENTIFIED SUBJECTS
2516	<input checked="" type="checkbox"/> LINK TEXT UNITS TO (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) LEARNING RESOURCES CONCERNING EACH TEXT UNITS

FIG. 25

26/69

LANGUAGE LEARNING SERVICES ASSOCIATED WITH LANGUAGE LEARNING PERSONALITY	
	SEL
2604	<input checked="" type="checkbox"/> SEL
2606	<input checked="" type="checkbox"/> LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR DEFINITIONS
2608	<input type="checkbox"/> LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR TRANSLATIONS
2610	<input type="checkbox"/> LINK EACH SENTENCE, OR PHRASE, TO A GRAMMATICAL DESCRIPTION OF THE STRUCTURE OF THE SENTENCE OR PHRASE
2612	<input type="checkbox"/> LINK EACH WORD, MULTWORD EXPRESSION, PHRASE OR SENTENCE TO OTHER INSTANCES OF THE SAME IN DIFFERENT CONTEXTS FROM THE PRESENT
2614	<input type="checkbox"/> LINK EACH WORD, MULTWORD EXPRESSION, PHRASE OR SENTENCE TO ONE OR MORE INTERACTIVE GRAMMAR EXERCISES CONCERNING THAT ELEMENT
2616	<input checked="" type="checkbox"/> LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO A LIST OF REAL WORLD CLASSES ABOUT IDENTIFIED LANGUAGE
	<input checked="" type="checkbox"/> LINK TEXT UNITS TO (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) LANGUAGE LEARNING RESOURCES CONCERNING EACH TEXT UNITS

FIG. 26

27/69

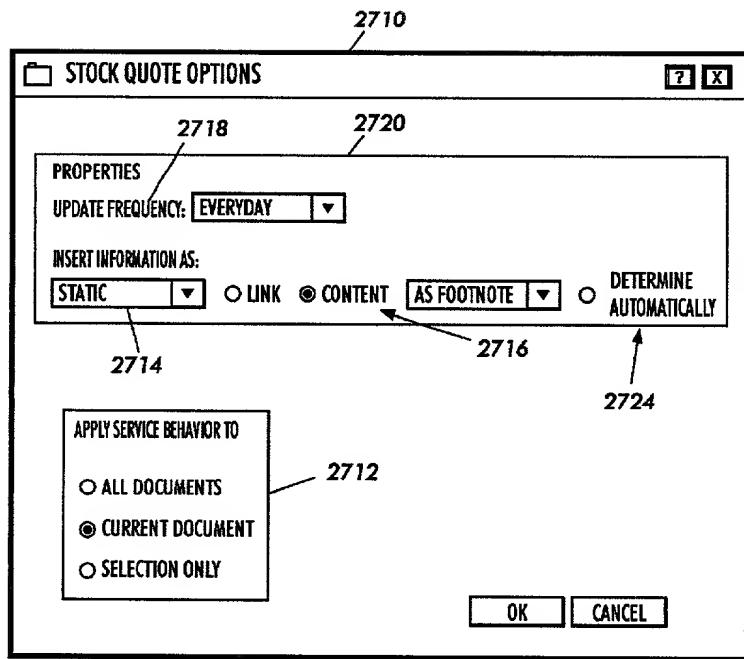


FIG. 27

28/69

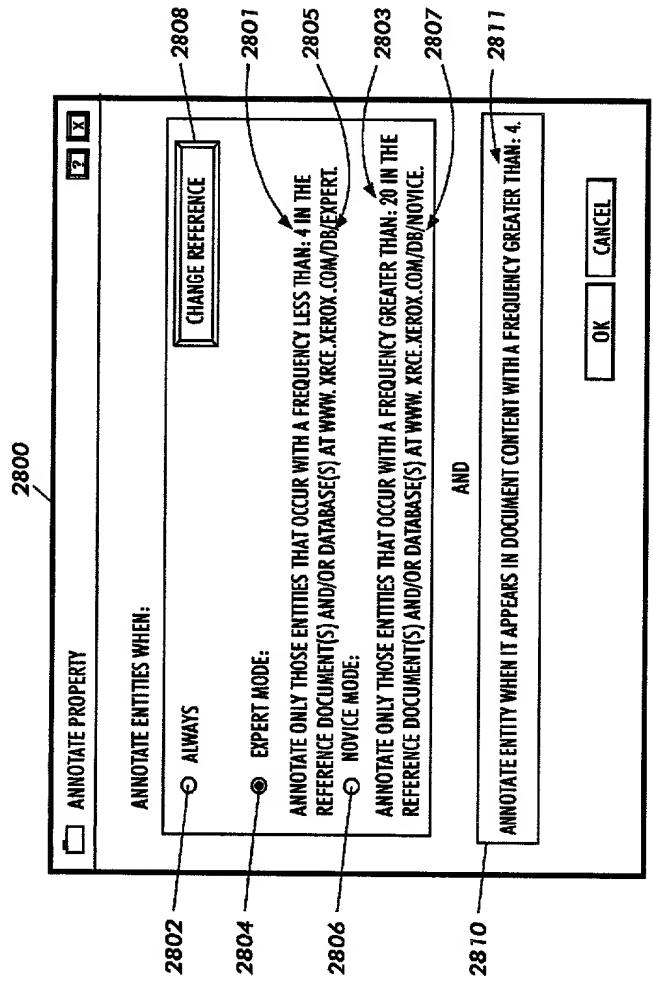


FIG. 28

29/69

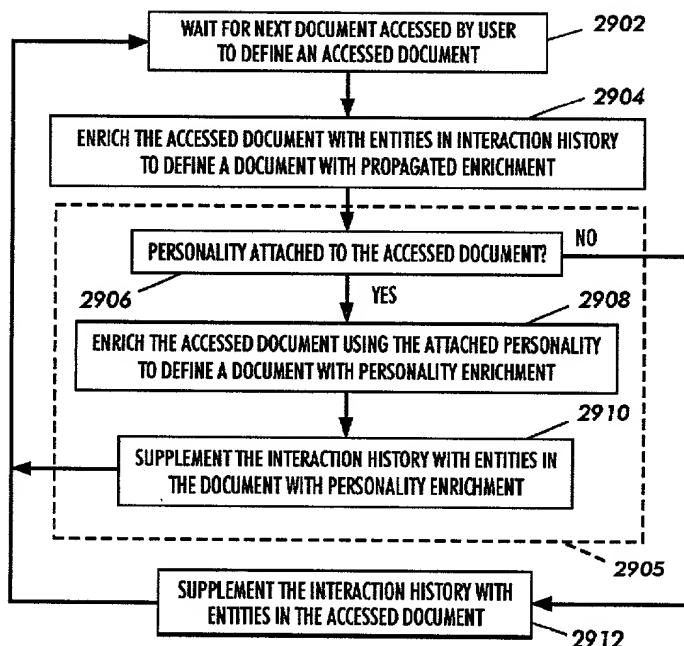


FIG. 29

30/69

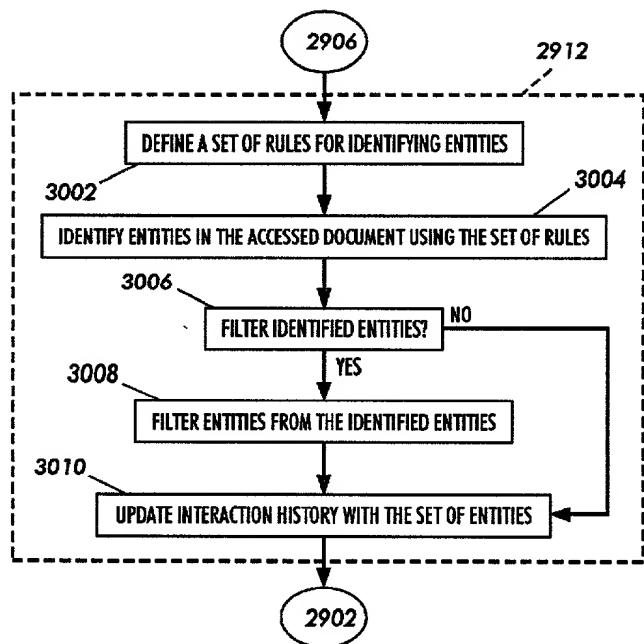


FIG. 30

31/69

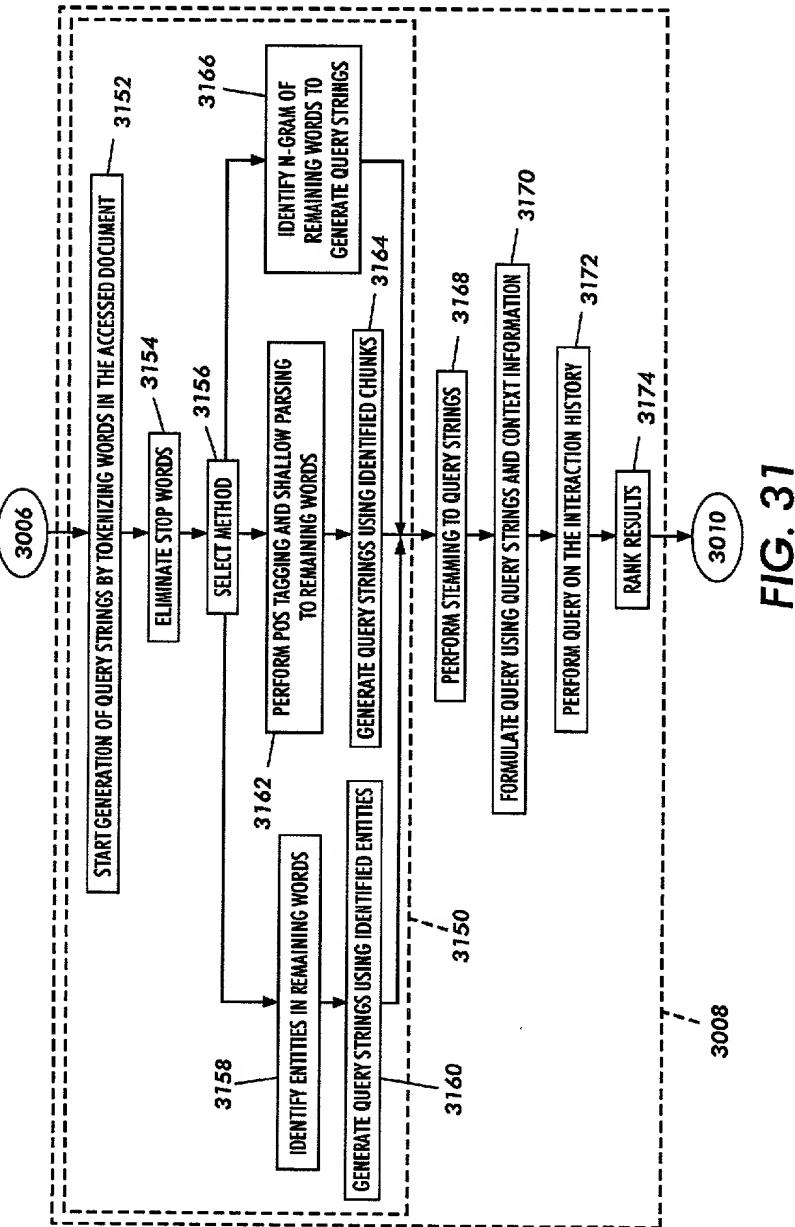


FIG. 31

32/69

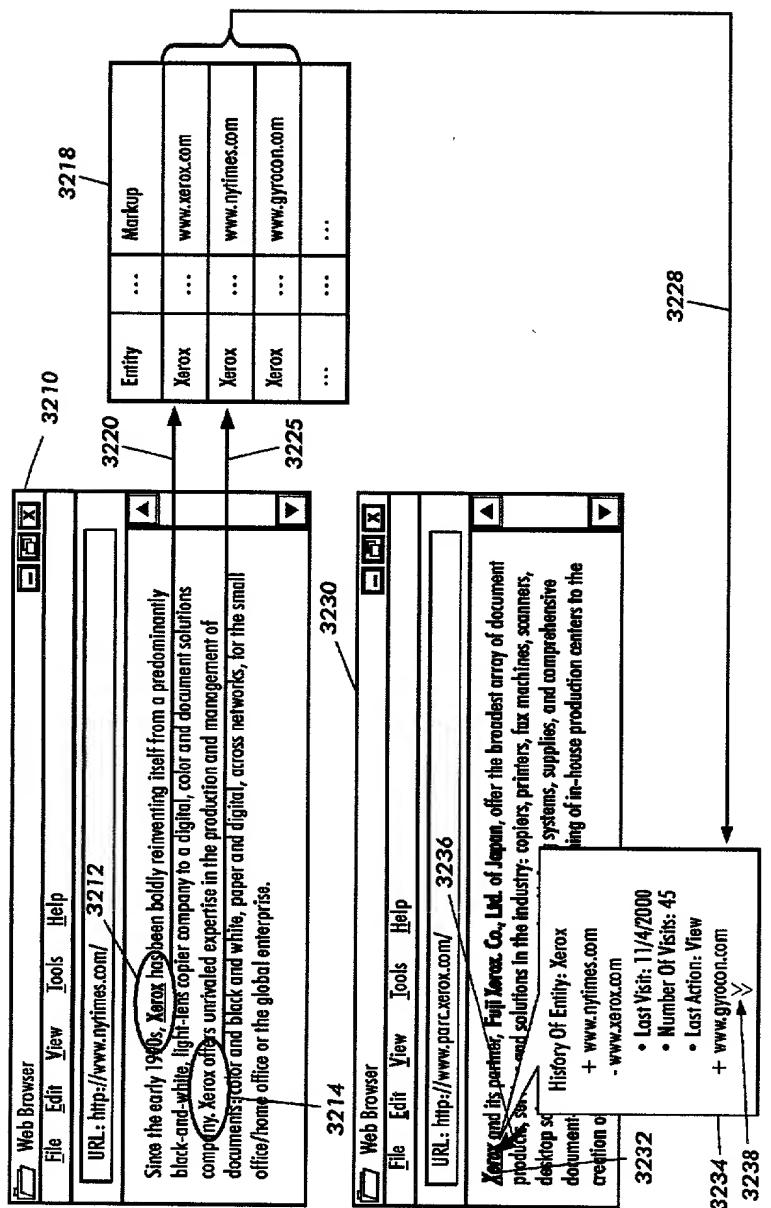


FIG. 32

33/69

3218

ENTITY	ENTITY TYPE	VISITS	LAST VISIT	ACTION LIST	DMOZ CATEGORY	LOCATION	POS	MARKUP
XEROX	BUSINESS	45	11/4/2000	VIEW(43), PRINT[2]	BUSINESS	40	NOUN	WWW.XEROX.COM
XEROX	BUSINESS	1	5/1/2001	VIEW(1)	FINANCE	689	NOUN	WWW.NYTIES.COM
XEROX	BUSINESS	4	2/22/2001	EMAIL(1), VIEW(3)	COMPUTER	343	NOUN	WWW.GYRO.COM.COM
...

FIG. 33

34/69

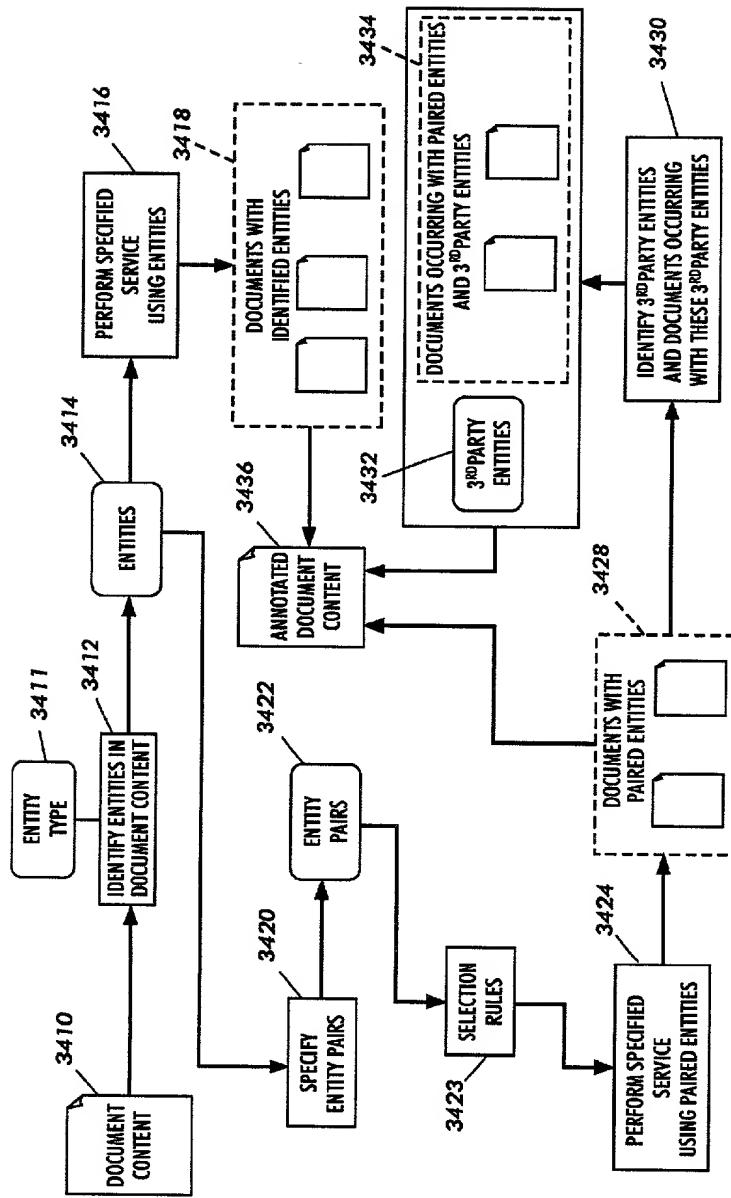


FIG. 34

35/69

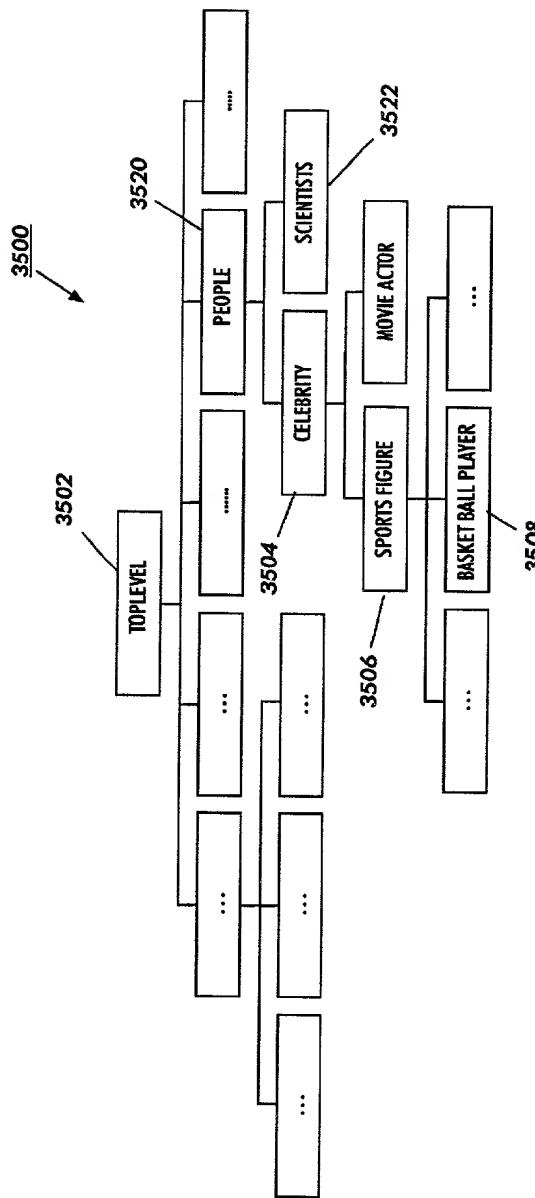


FIG. 35

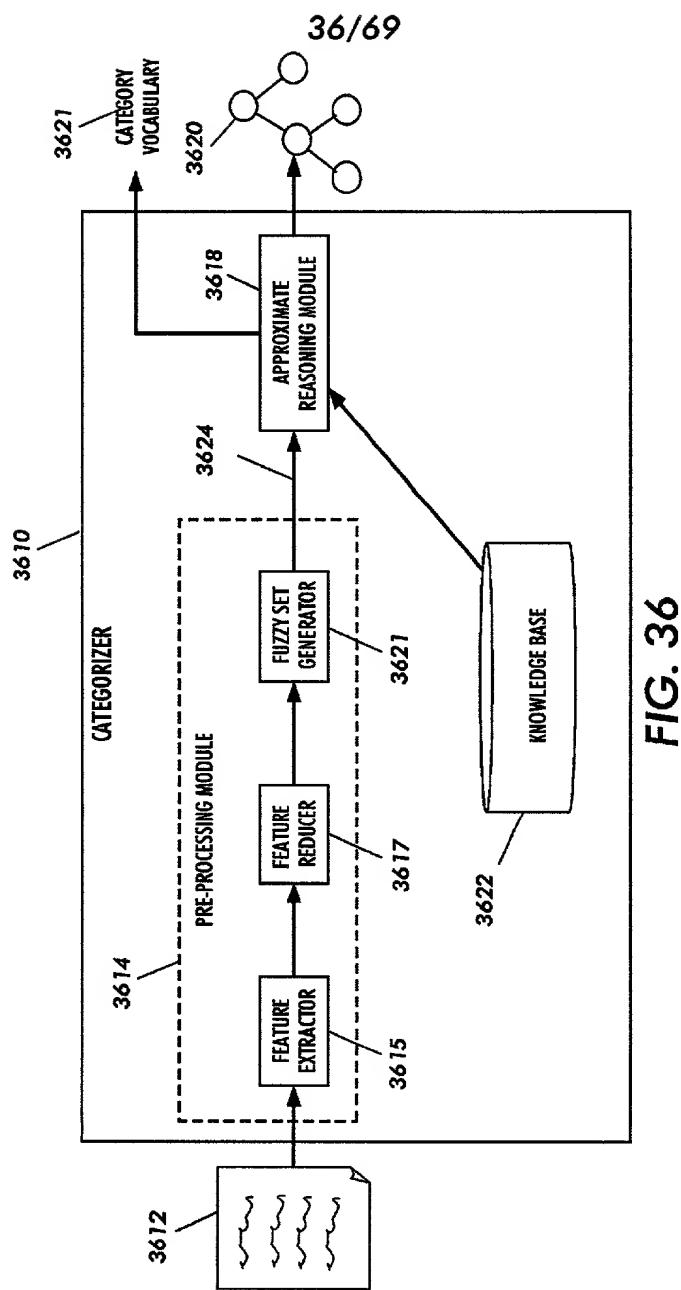


FIG. 36

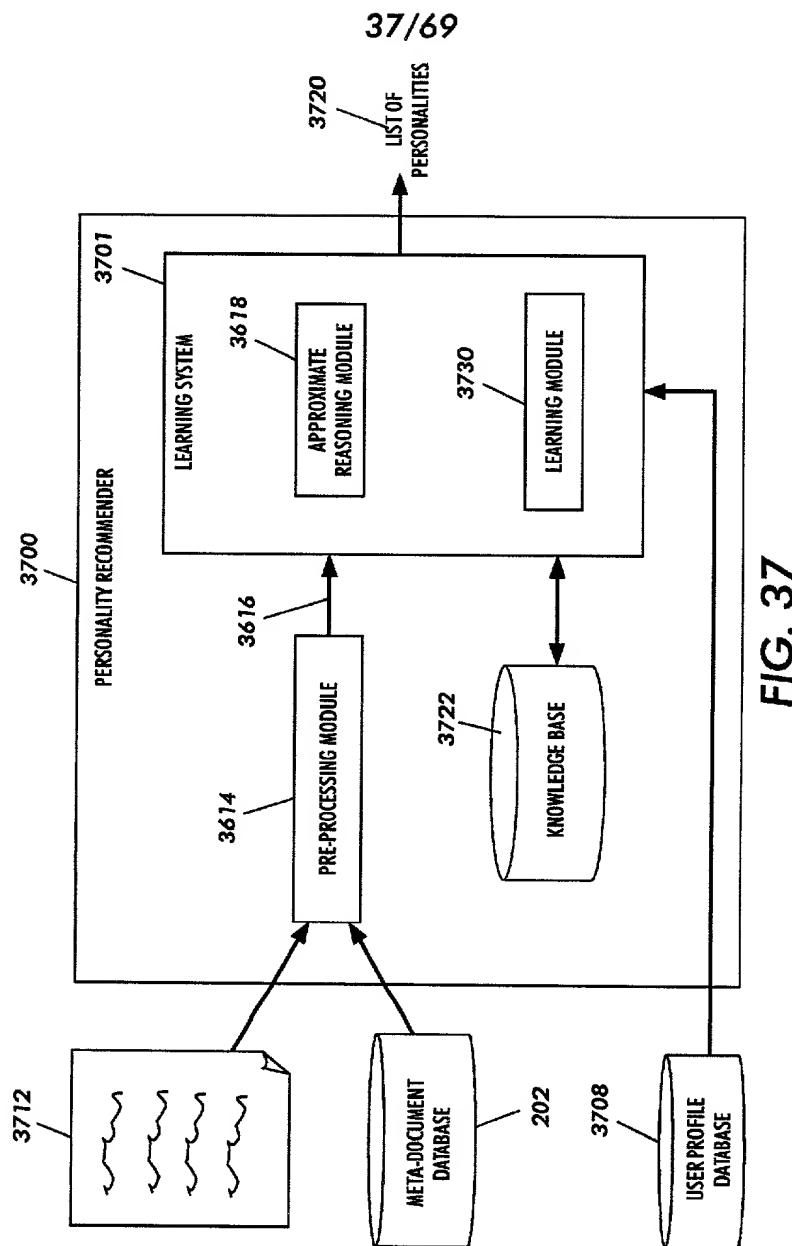


FIG. 37

38/69

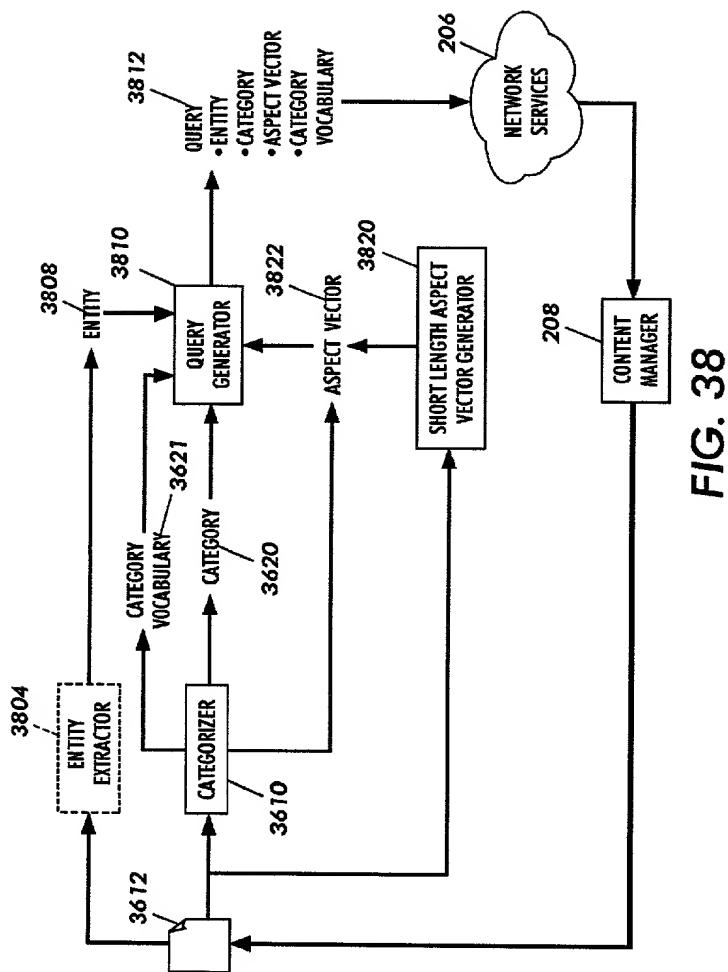


FIG. 38

39/69

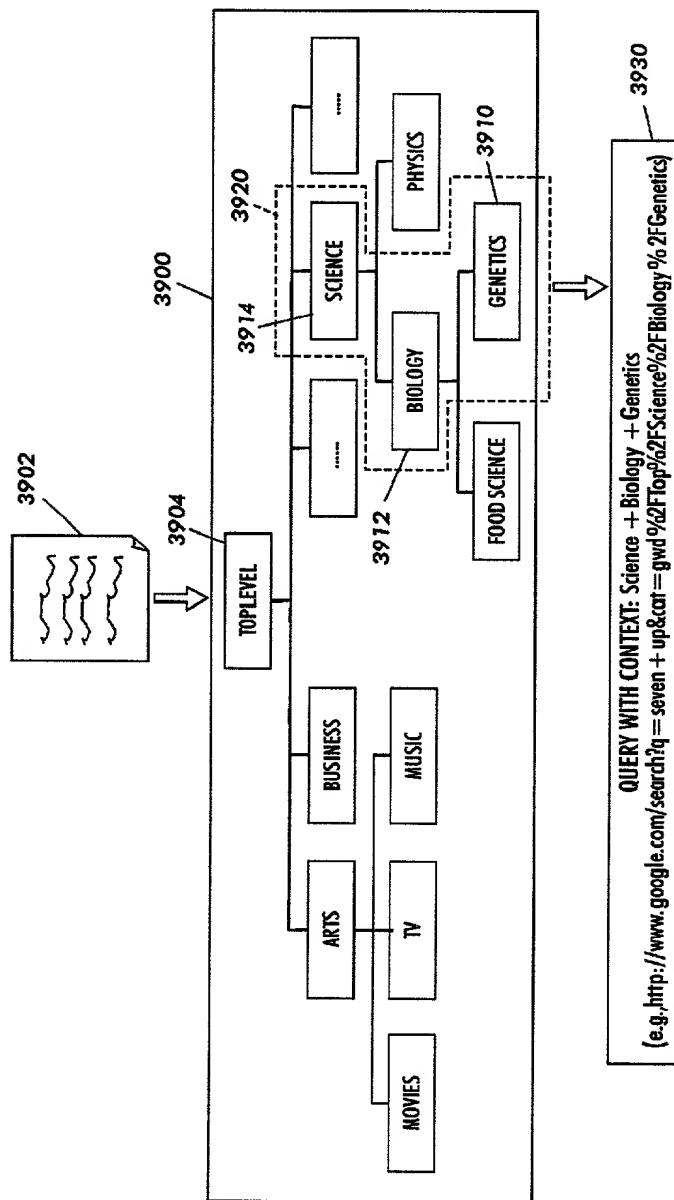


FIG. 39

40/69

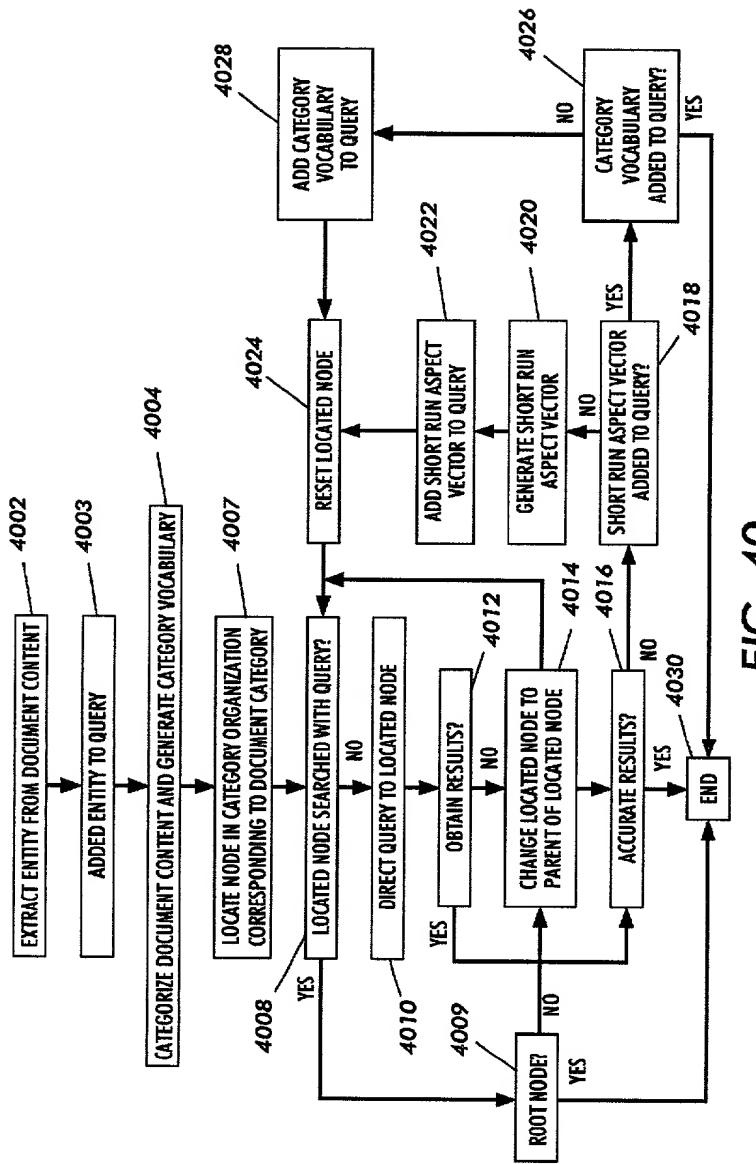
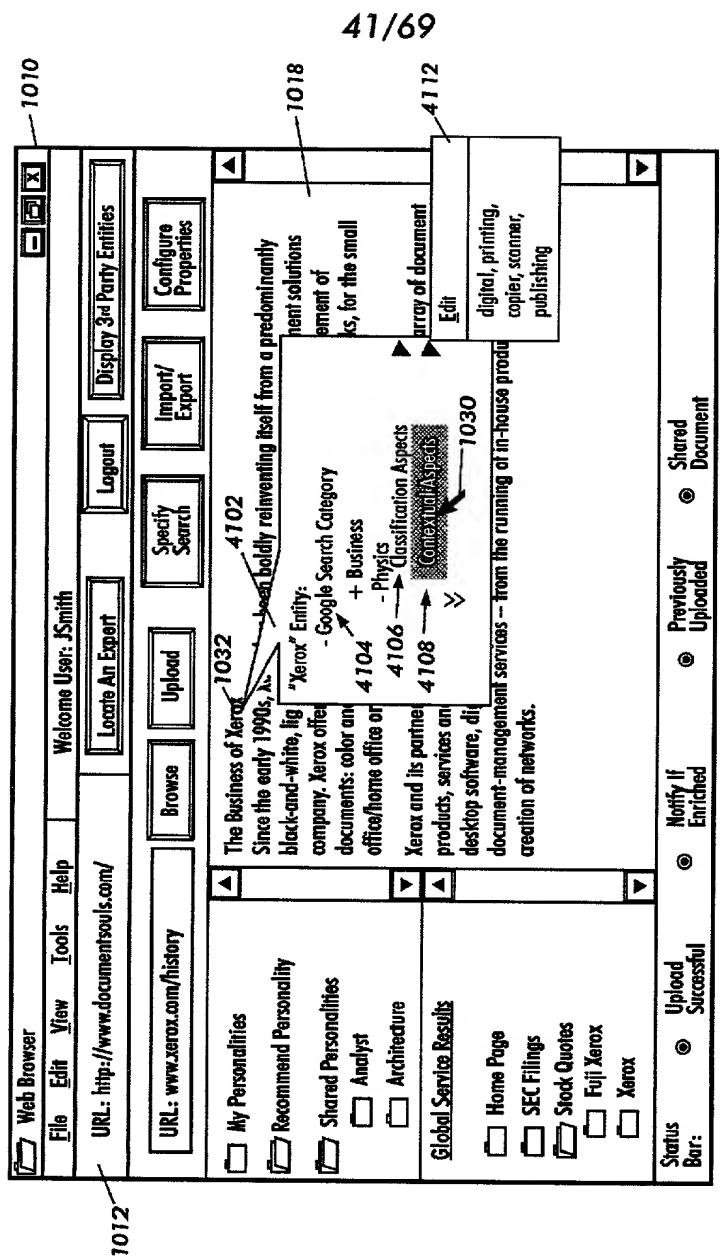
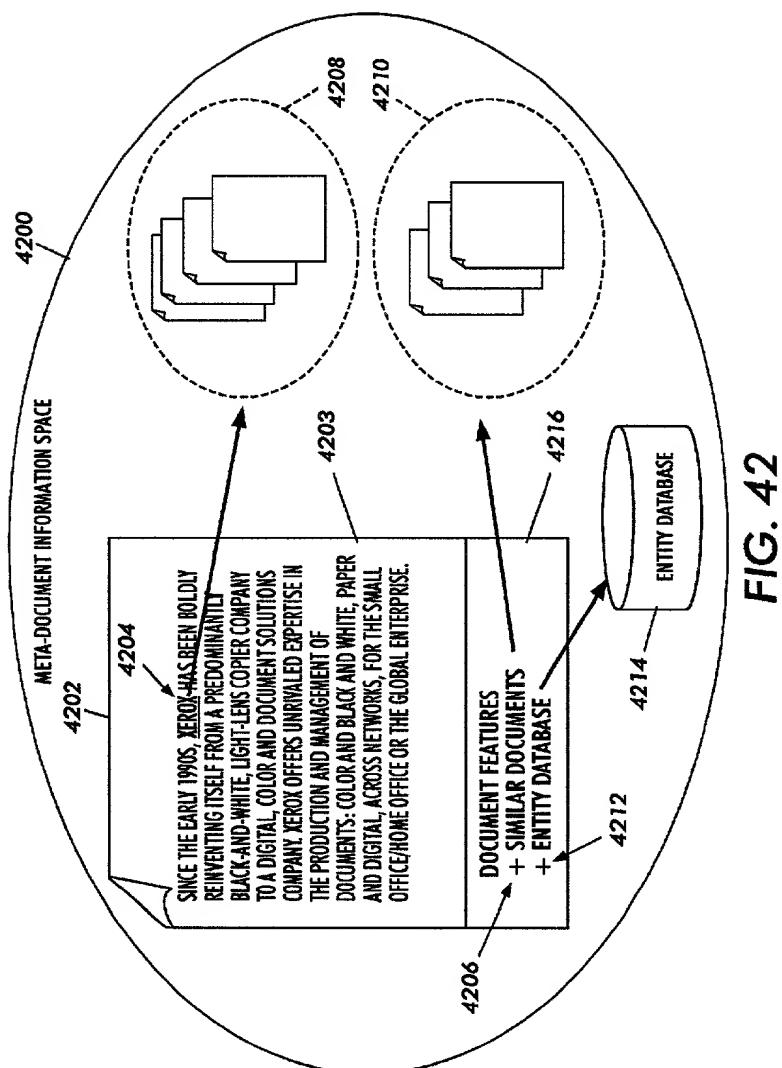


FIG. 40



42/69



43/69

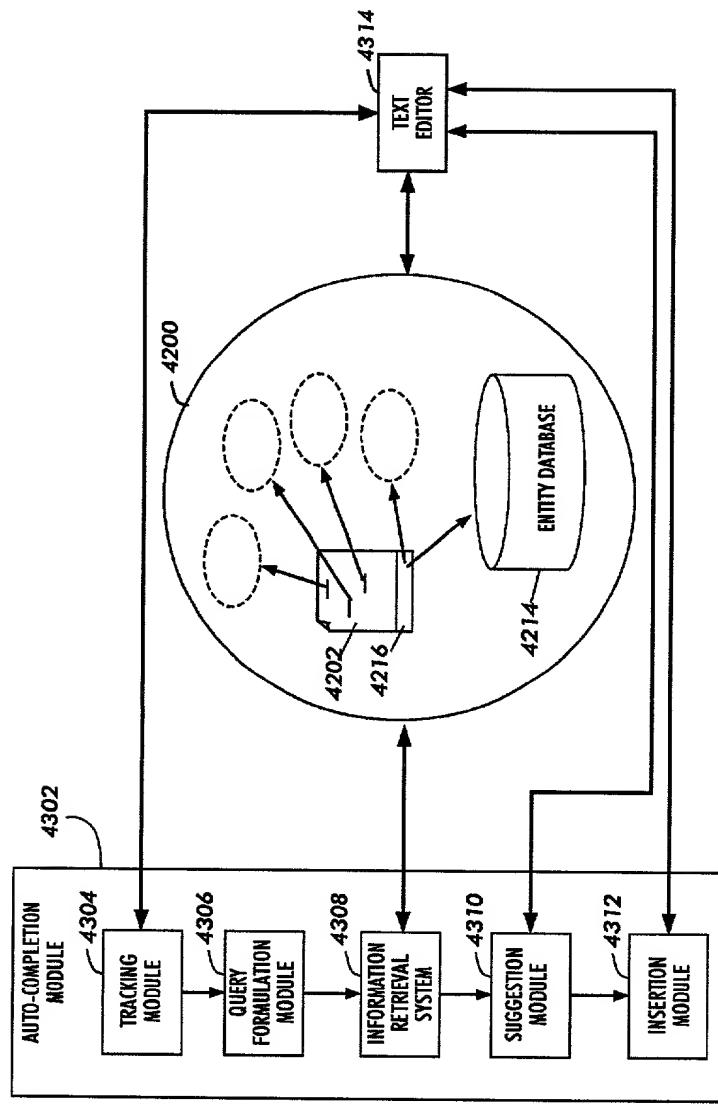


FIG. 43

44/67

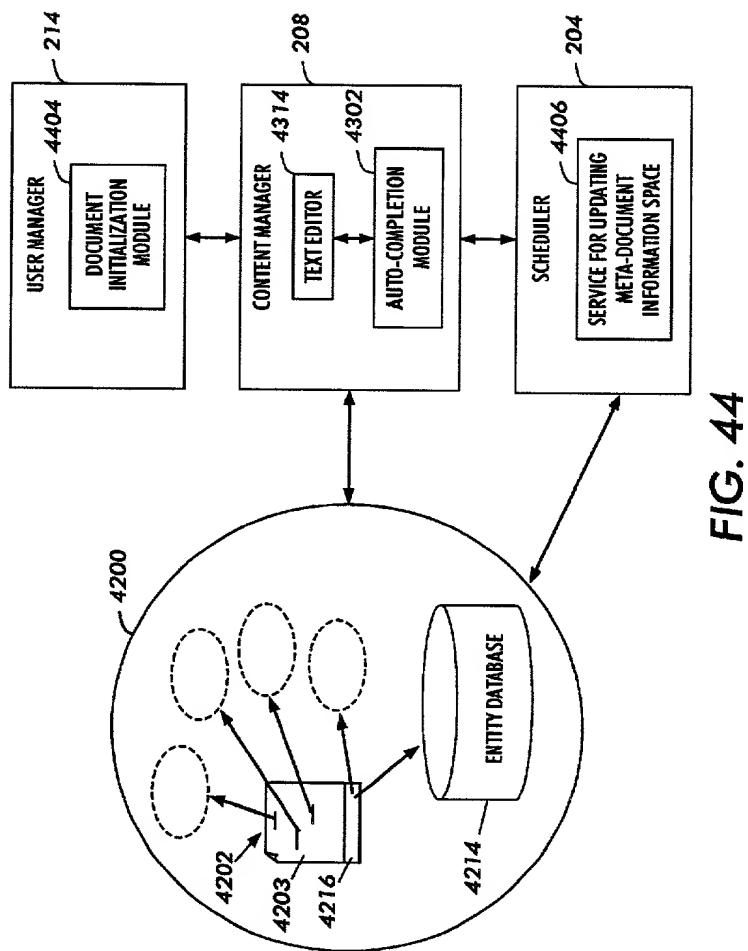


FIG. 44

45/67

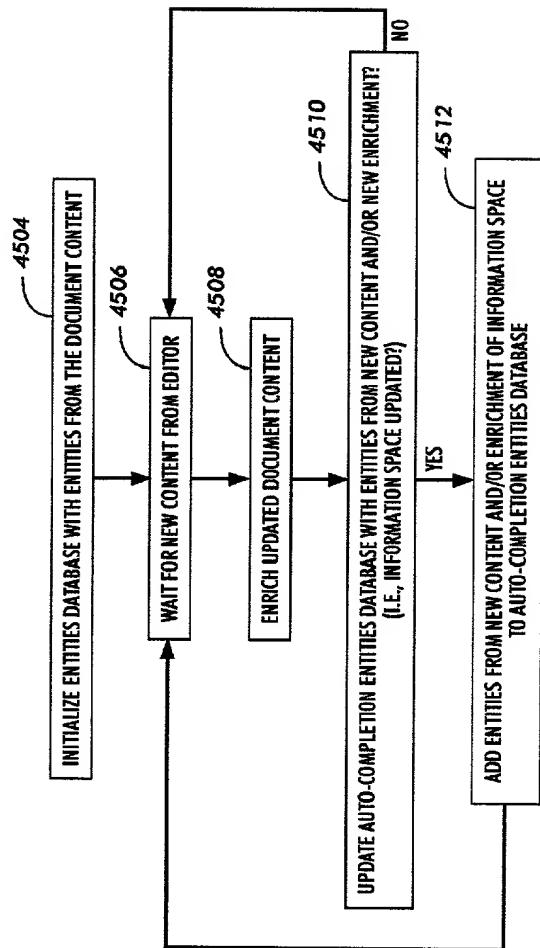


FIG. 45

46/67

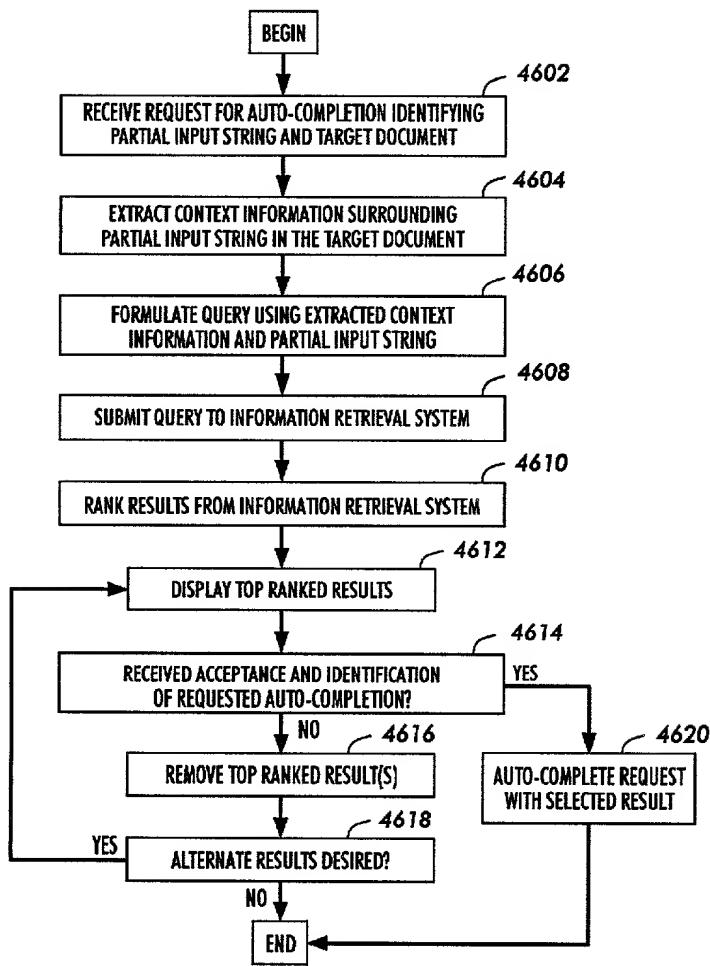


FIG. 46

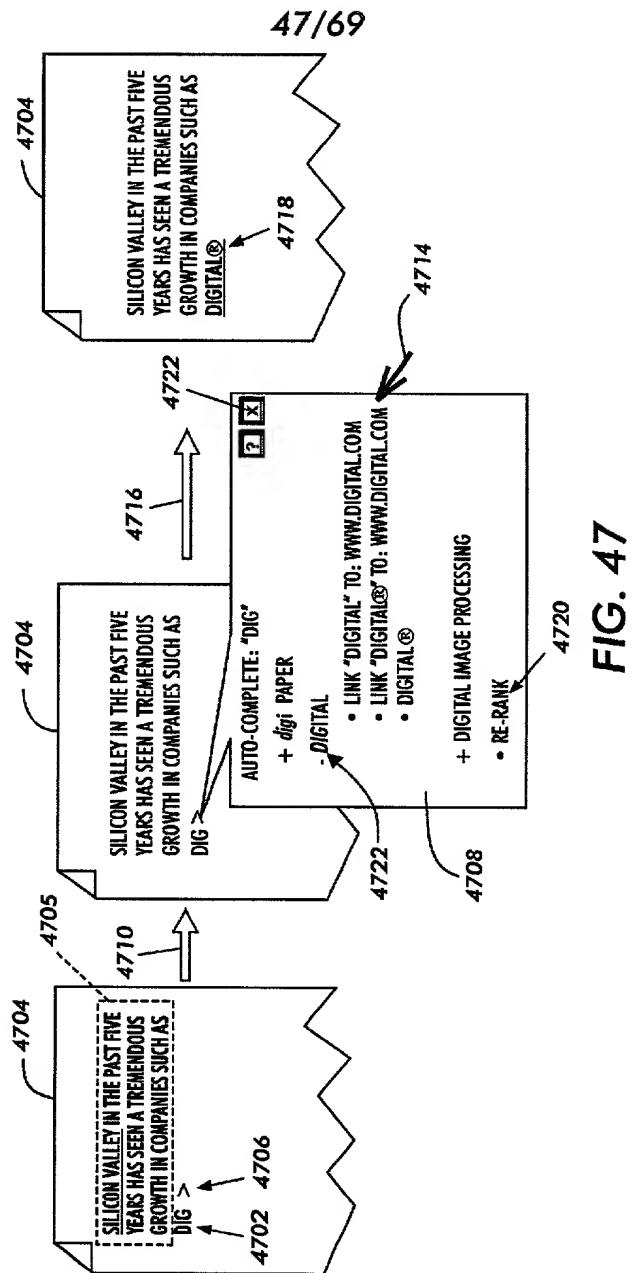


FIG. 47

48/69

(A) KEY	(B) ENTITY	(C) MARKUP	(D) FORMAT	(E) ORIGIN CONTENT	(F) ORIGIN POSITION	(G) PART OF SPEECH	(H) CONTEXT
Sili	Silicon Valley	www.siliconvalley.com	Capitalize	TARGET DOCUMENT	1	Noun	Science
Dig	Digital	www.digital.com	®	www.digital.com	424	Adjective	Science
Dig	DigiPaper	NONE	Bold	www.parc.xerox.com	23	Noun	Science
Dig	digress	NONE	NONE	www.nytimes.com	343	Verb	Politics
...

FIG. 48

49/69

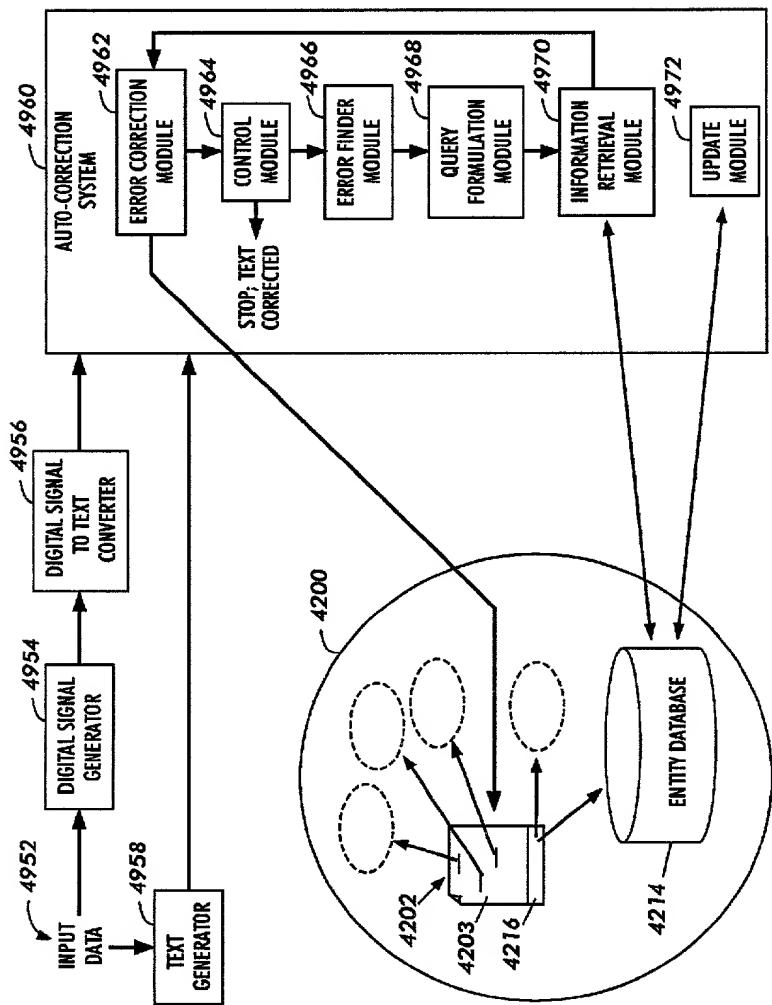


FIG. 49

50/69

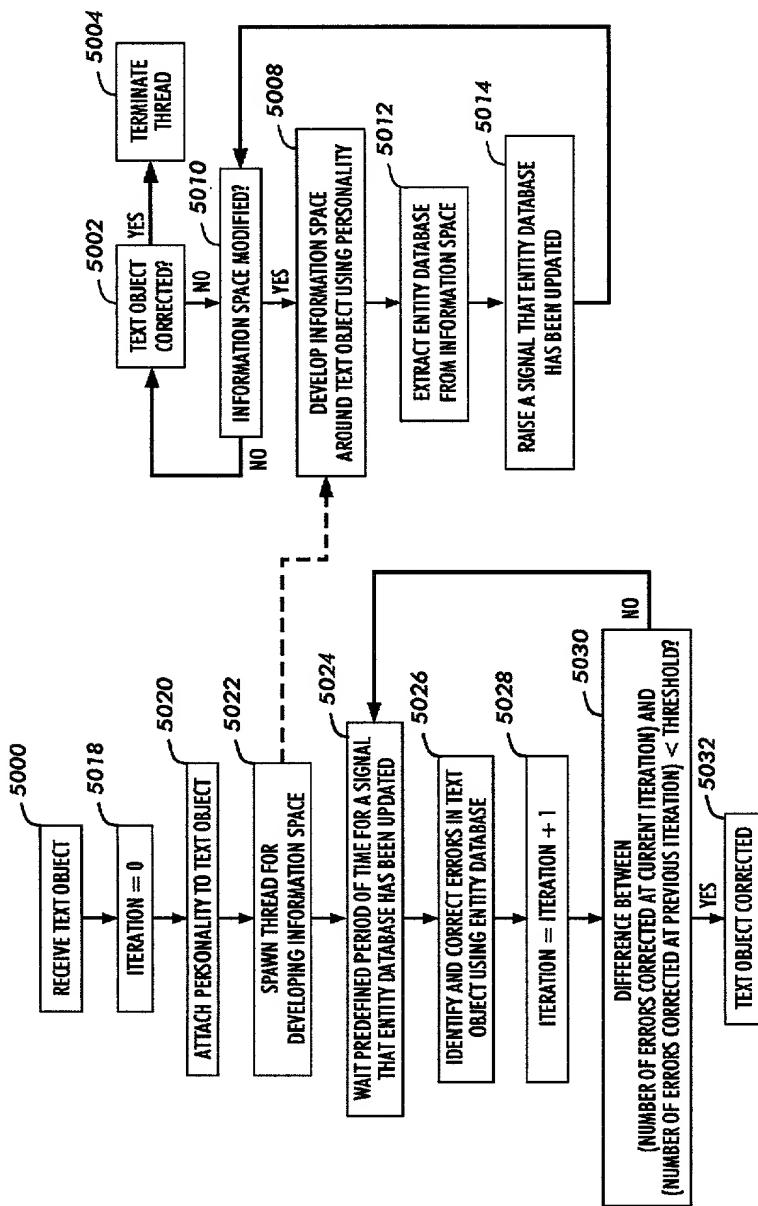
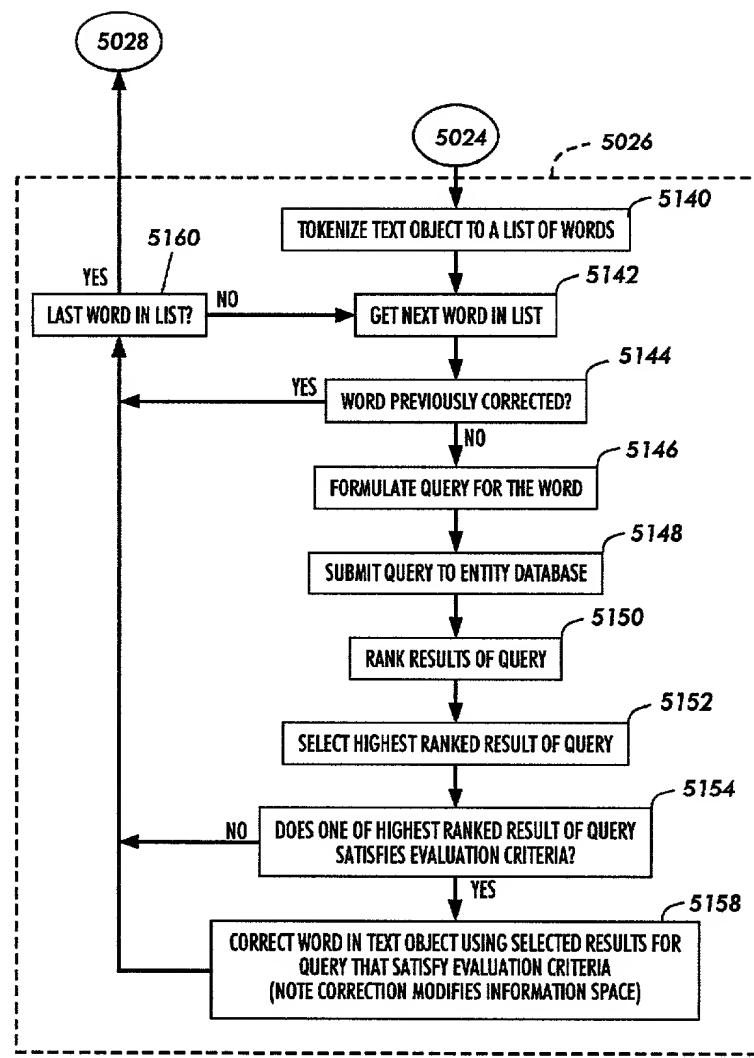


FIG. 50

51/69

**FIG. 51**

52/69

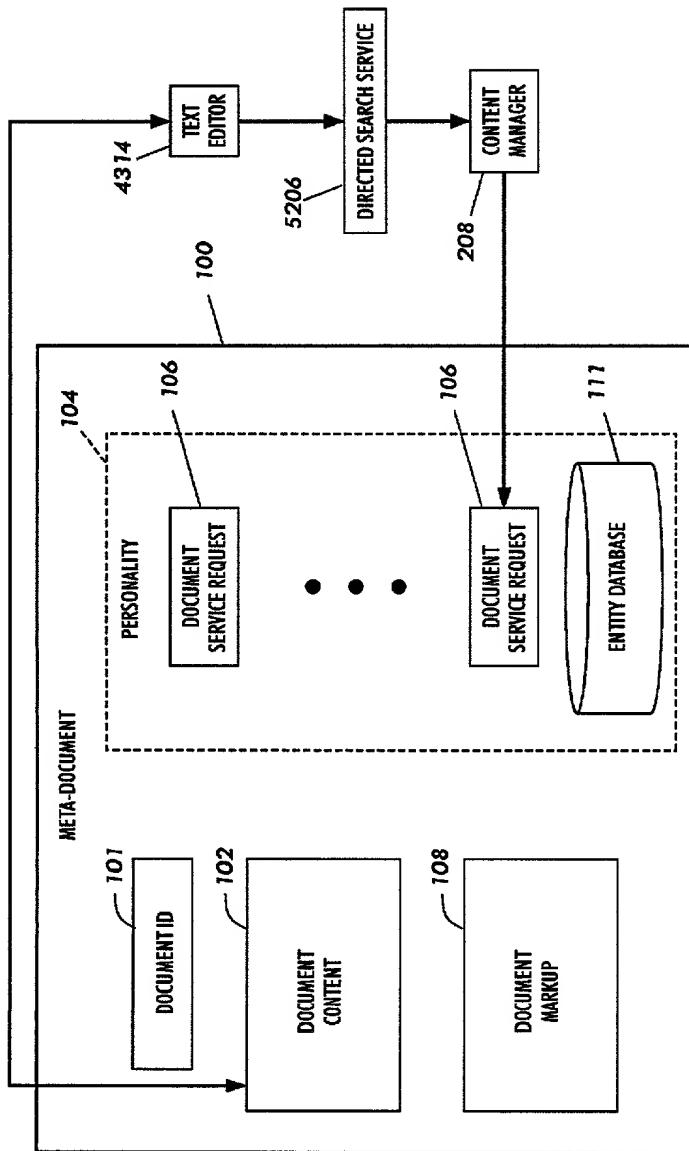


FIG. 52

53/69

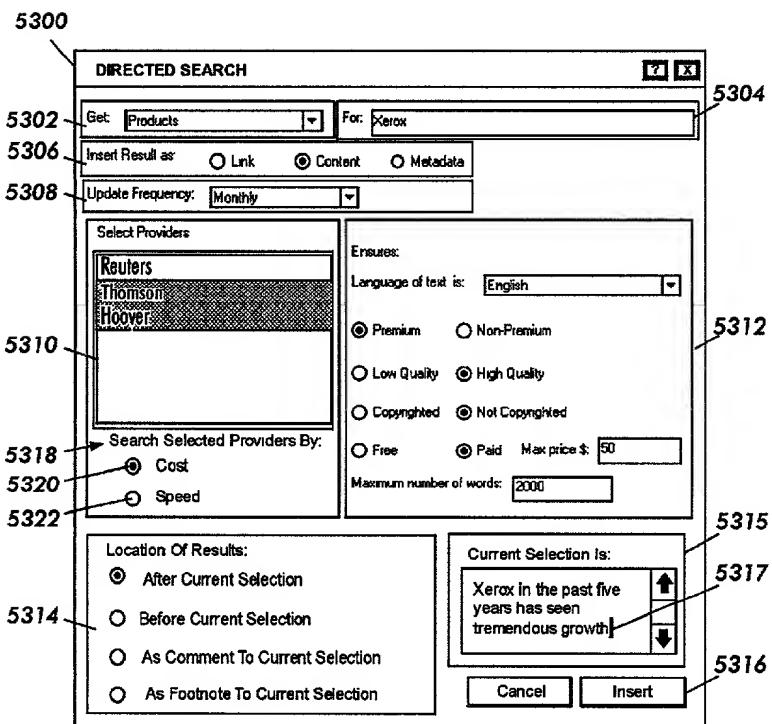


FIG. 53

54/69

5400

```
<service id='Products'

    refresh=30
    insert='Content'
    language='En'
    quality='Premium&High&Not-copyrighted&paid'
    maxprice='50'
    maxwords='2000'
    search='cost'
    <entity>Xerox
    <offsets>13 18 456 461 710 715</offsets>
    <provider name='Thomson' id=18></provider>
    <provider name='Reuters' id=19></provider>
    </entity>
  </service>
```

FIG. 54

55/69

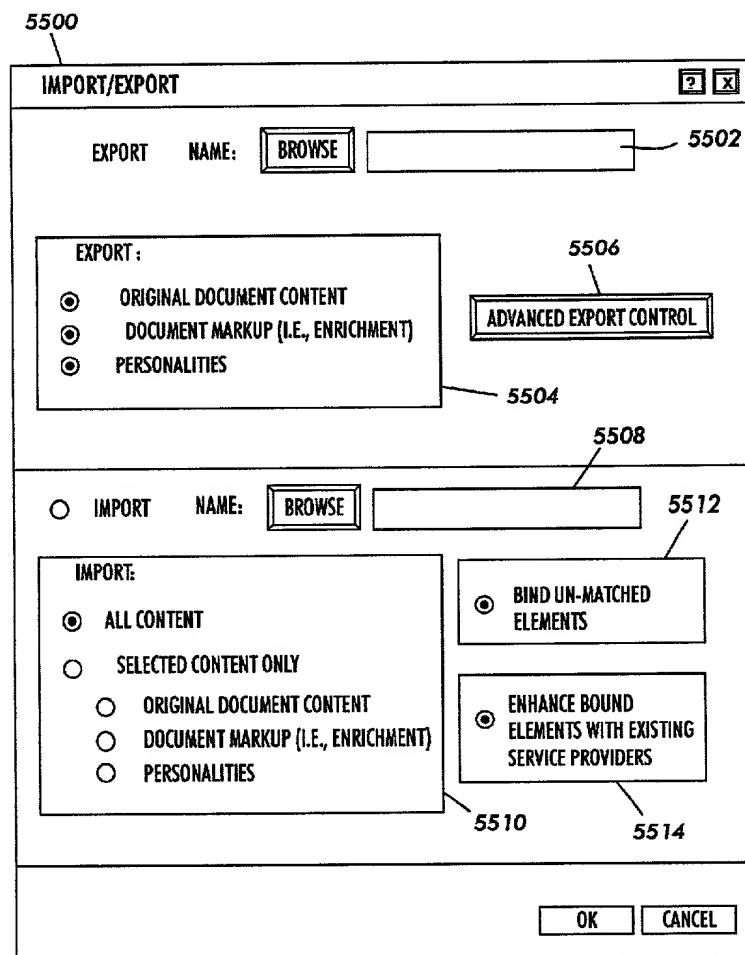


FIG. 55

56/69

```

<!-- SERVER ID>
<docSouls:serverId> DS v1.0 server name </docSouls:serverId>
<!-- PROPERTIES OF DOCUMENT>
<docSouls:docId> xxx </docSouls:docId>
<docSouls:name> xxx </docSouls:name>
<docSouls:owner> xxx </docSouls:owner>
''

<!-- ORIGINAL DOCUMENT CONTENT>
<docSouls:originalDoc format="HTML">toto.html</docSouls:originalDoc>
<docSouls:browsingDoc format="HTML">browse.html</docSouls:browsingDoc>

<!-- LIST OF PERSONALITIES AND DOCUMENT SERVICE REQUESTS>
<docSouls:endowmentDescription>
''

<!-- LIST OF PERSONALITIES>
<docSouls:personality id = " " name="" etc.>
<docSouls:service id="" name="" etc.>
<docSouls:provider name="" etc.>
<docSouls:provider name="" etc.>
</docSouls:service>
<docSouls:personality>
''

<!-- LIST OF SERVICE REQUESTS>
<docSouls:service id="" name="" etc.>
<docSouls:provider name="" etc.>
<docSouls:provider name="" etc.>
</docSouls:service>
''

</docSouls:endowmentDescription>

```

FIG. 56

57/69

```

<!-- DESCRIPTIONS OF PERSONALITIES>
<docsouls:personalitydescription>
<docsouls:id> xx </docsouls:id>
<!-- all the fields AND service categories
<docsouls:servicecategory id="" name = "">
</docsouls: personalitydescription>

<!-- NEXT PERSONALITY DESCRIPTION>
...
<!-- DESCRIPTION OF DOCUMENT SERVICE REQUESTS>
<docsouls:servicedescription>
<docsouls:id> xx </docsouls:id>
<!-- list all fields (e.g., "refresh period", category, etc.)>
<!-- list all providers>
<docsouls:provider id="" name = "">
<!-- list all strategies>
<!-- list status of services (original, frozen, mapped, etc.)>
<docsouls:servicestatus status="mapped" originalid="" originalserver="">
<!-- list ServerId if the service was not available on current server>
<docsouls:serverid></docsouls:serverid>
</docsouls:servicedescription>

<!-- NEXT SERVICE REQUESTS DESCRIPTION>
...

```

FIG. 57

58/69

```

<!-- DESCRIPTION OF PROVIDERS >
<docsouls:providerdescription>
  <docsouls:id> xx </docsouls:id>
    <!-- all the fields e.g. key, URI, etc. >
  </docsouls: providerdescription >

<!-- DESCRIPTION OF DICTIONARIES >
<docsouls:dictionarydescription>
  <docsouls:id> xx </docsouls:id>
    <!-- all the fields, e.g. category >
  </docsouls:dictionarydescription >

<!-- DESCRIPTION OF STRATEGY >
<docsouls:strategydescription>
  <docsouls:id> xx </docsouls:id>
    <!-- all the fields >
    <!-- all the dictionaries >
  </docsouls: strategydescription >

<!-- RESULTS OF GLOBAL SERVICES >
<docsouls:globalservice id="Summary">
  <docsouls:link> xxx </docsouls:link>
  <docsouls:content> xxx </docsouls:content>
</docsouls:globalservice>

<!-- OTHER RESULTS (E.G. KEYWORDS, CATEGORIES, ETC. >

```

FIG. 58A

59/69

```

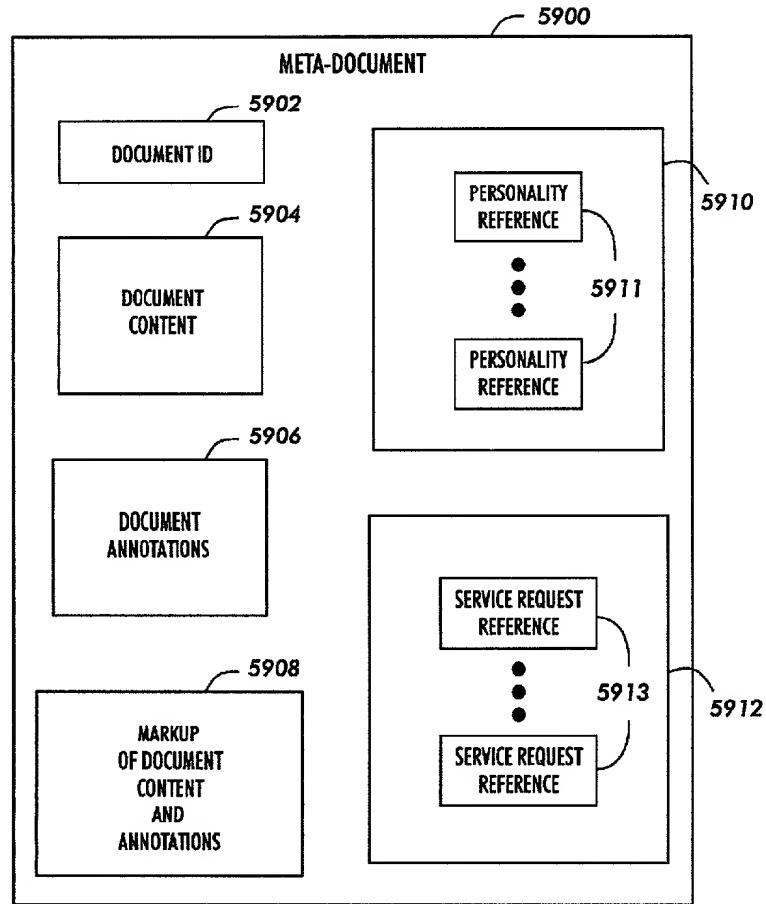
<!-- (OPTIONAL) DESCRIPTION OF ENRICHMENT PERFORMED AT OTHER DOCUMENT ENRICHMENT
SYSTEMS>
<docsouls:serverId>olderServerId</docsouls:serverId>
<docsouls:personalityDescription>...</docsouls:personalityDescription>
<docsouls:serviceDescription>...</docsouls:serviceDescription>

<!-- (OPTIONAL) LIST OF RESULTS OF THE LAST ENRICHMENT>
<!-- IDENTIFY DOCUMENT SECTION IN CASE OF MULTI-PART DOC >
<docsouls:docpart>main</docsouls:docpart>
<!-- OFFSET WHEN ENRICHMENT APPLIES TO A SECTION >
<docsouls:offset> 0 200 </docsouls:offset>
<!-- STATUS OF SERVICE APPLIED ("ACTIVE" OR "FROZEN") >
<docsouls:service id="StockQuote" status="active">
  <!-- ENTITY>
    <docsouls:entity>xerox</docsouls:entity>
  <!-- NORMALIZED FORM >
    <docsouls:normForm>xerox</docsouls:normForm>
  <!-- LIST OF OFFSETS WHERE ENTITY IS FOUND IN TEXT RELATIVE TO SECTION >
    <docsouls:offsets start=35 end=40 />
    <docsouls:offsets start=232 end=237 />
  <!-- DICTIONARIES >
    <docsouls:dict> xxx </docsouls:dict>
  <!-- STRATEGIES >
    <docsouls:strat> xxx </docsouls:strat>
  <!-- LINK INFORMATION >
    <docsouls:link> xxx </docsouls:link>
  <!-- CONTENT INFORMATION >
    <docsouls:content> xxx </docsouls:content>
  <!-- NEXT ENTITY>
    ...
</docsouls:service>

```

FIG. 58B

60/69

**FIG. 59**

61/69

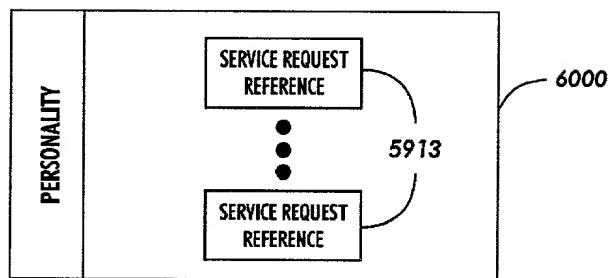


FIG. 60

62/69

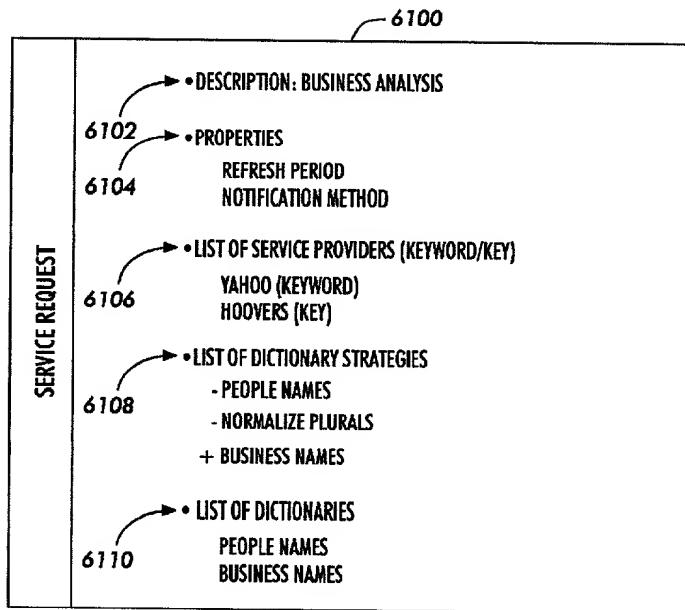


FIG. 61

63/69

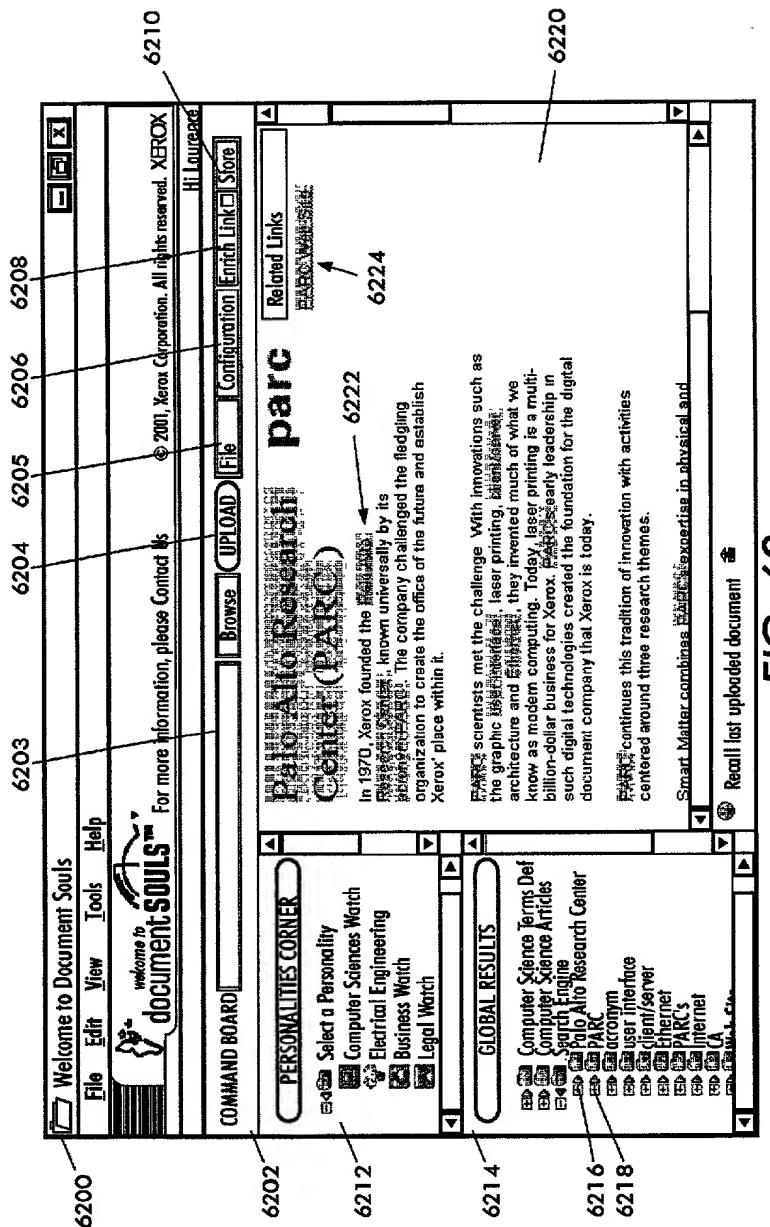


FIG. 62

64/69

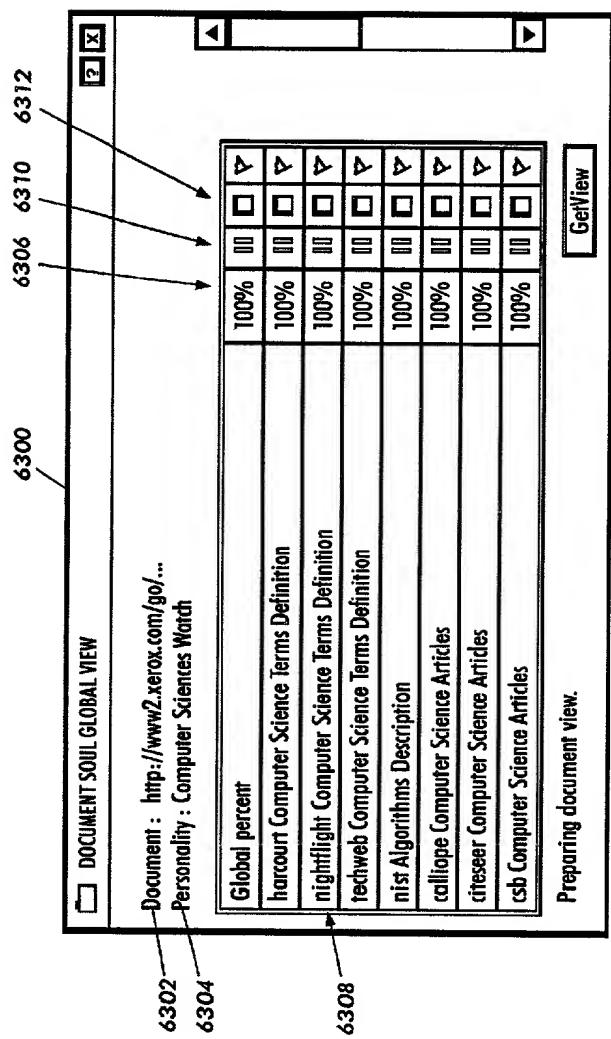


FIG. 63

65/69

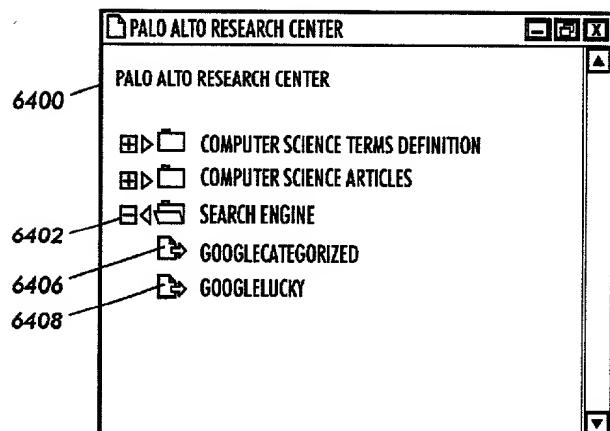


FIG. 64

66/69

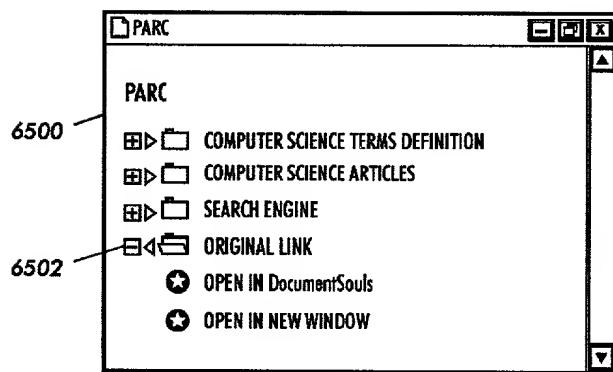


FIG. 65

67/69

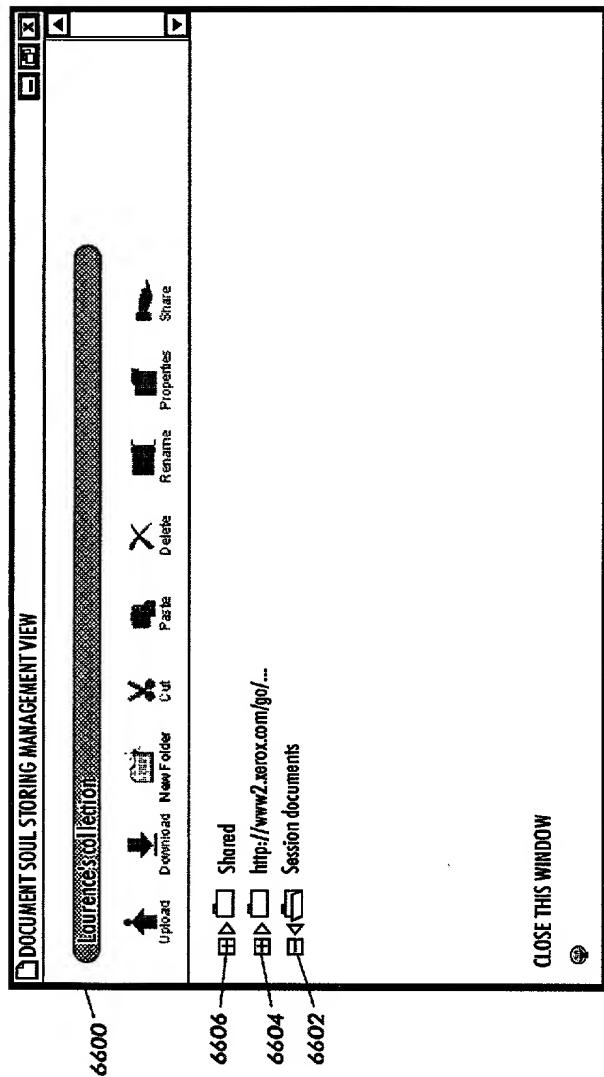


FIG. 66

68/69

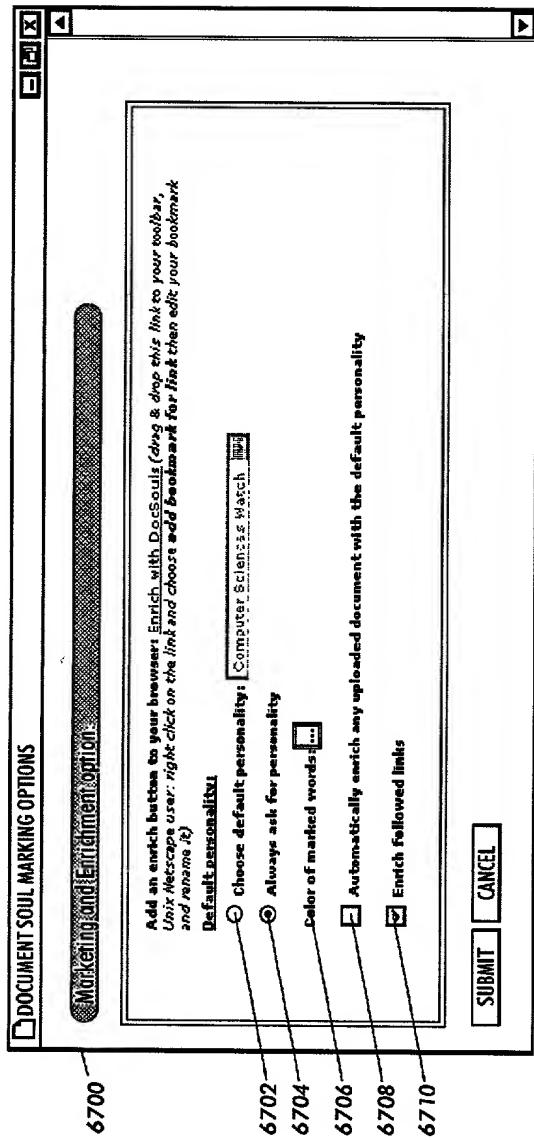


FIG. 67

69/69

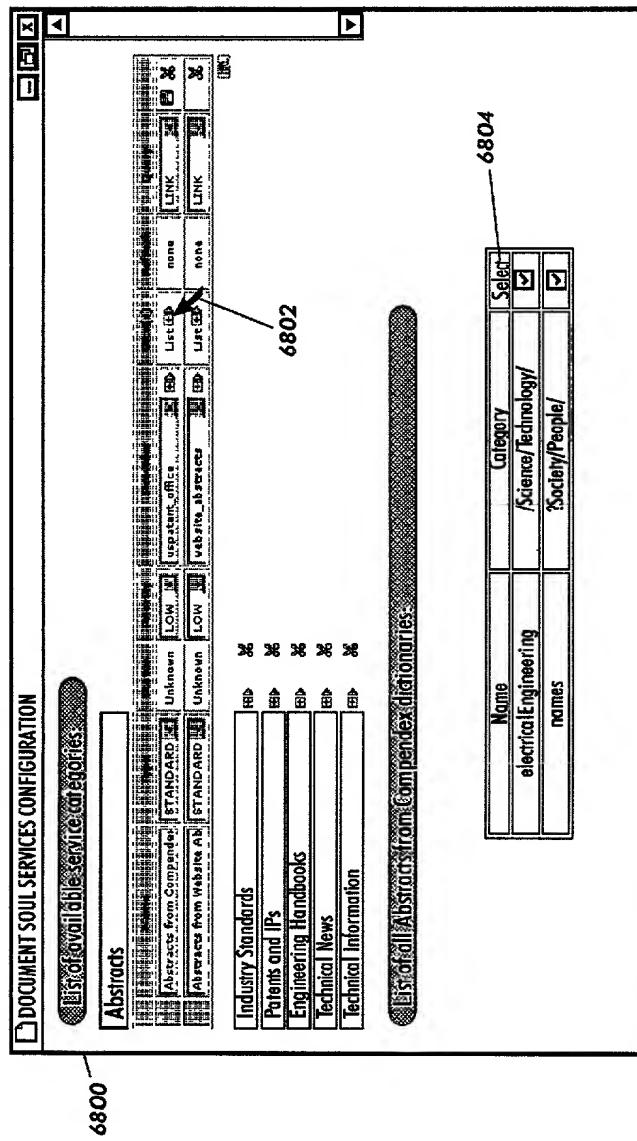


FIG. 68